SEP 29 1922

VOL. V, NO. 4 PSYCHOLOGICAL REVIEW PUBLICATIONS AUGUST, 1922

Journal of Experimental Psychology

EDITED BY

JOHN B. WATSON, 244 MADISON AVE., NEW YORK CITY
HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
JAMES R. ANGELL, YALE UNIVERSITY (*Monographs*)
SHEPHERD I. FRANZ, GOVT. HOSP. FOR INSANE (*Bulletin*) AND
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*Index*)

ADVISORY EDITORS

H. A. CARR, UNIVERSITY OF COLORADO; RAYMOND DODGE, WESLEYAN UNIVERSITY; KNIGHT DUNLAP, THE JOHNS HOPKINS UNIVERSITY; H. S. LANGFELD, HARVARD UNIVERSITY; E. L. THORNDIKE, COLUMBIA UNIVERSITY.

CONTENTS

Individual Differences in Retinal Sensitivity, and their Correlation with [illegible] *Findings:* PERCY W. COBB 227

Additional Tests of the Stanford Revision Applied to University Faculty Members: HERBERT CALDWELL 247

[illegible] J. CLEMENT CHAPMAN 263

[illegible] Frequency and Tendency Factors are Negative: JOSEPH PETERSON 270

PUBLISHED BIMONTHLY BY
PSYCHOLOGICAL REVIEW COMPANY
[illegible] AND LANCASTER STREETS, LANCASTER, PA., AND
PRINCETON, N. J.

Entered, July 1, 1920, at the post office at Lancaster, Pa.

Journal of
Experimental Psychology

VOL. V., NO. 4 AUGUST, 1922

INDIVIDUAL VARIATIONS IN RETINAL SENSITIVITY, AND THEIR CORRELATION WITH OPHTHALMOLOGIC FINDINGS[1]

BY CAPTAIN PERCY W. COBB, M.C.

Medical Research Laboratory, Air Service, U. S. Army, Mitchel Field, Garden City, L. I., N. Y.

ACKNOWLEDGMENT

The author wishes to acknowledge the help of Dr. Mildred W. Loring and Miss Elizabeth K. Stark in the performance of most of the technical work of accumulating these results; the help of Miss Marion L. Smith in performing the computations; and the courtesy of Nela Research Laboratories in facilitating the completion of the manuscript.

The method by which the measurements of retinal sensitivity were made, which are the subject of this communication, has, in its essential details, already been described.[2] The purpose of the present paper is to investigate primarily the significance of the results obtained by the method, and secondarily its applicability as a test in the routine examination of candidates for flying status in the Air Service. Measurements were made upon a group of 101 subjects. Visual acuity was measured in nearly all of these, for each eye separately and for both eyes together by means of the Snellen letters under the actual conditions of the experiment; and in addition, comparison was made with the complete findings of the routine ophthalmological examination for flying status upon 59 of these subjects.

The group consisted of such individuals as were available,

[1] Published with permission of the Surgeon General, U. S. Army, who is not responsible for any opinion expressed or conclusions reached herein.

[2] Cobb and Loring, JOUR. EXP. PSYCH., Vol. IV., 1921, pp. 175 to 197.

and without specific enumeration, may be described by noting the classes into which they might be divided, in the order of their relative numbers: (1) Enlisted men of the army, a few from the navy, some recruits and some civil candidates for the cadet-school of aviation; (2) a number of flying officers; (3) a number of other officers, chiefly of the medical corps, of the staff of the Medical Research Laboratory or there present for instruction; (4) a few civil employes of the Laboratory (including four women), technical, clinical and one skilled mechanic.

The present method of estimating retinal sensitivity is, essentially, to determine the duration of exposure of a black dot at the center of a white screen, necessary in order that the dot may just be seen. The actual dimensions of the dot, and the distance of the subject from the screen are kept constant, as is also the photometric brightness of the screen. The dot is formed by a small circular aperture in the screen, through which a secondary screen, behind the other, is seen; and the screen behind is so illuminated that its appearance is an exact match with the margins of the aperture. When this is accomplished the aperture disappears. It becomes in appearance a black dot when an opaque object cuts off the light coming through it, and the necessary time of darkening, in order that the dot may be perceived with 50 per cent. certainty, is the quantity measured.

The usefulness of a new test may depend upon one of two circumstances. It may be supplementary to tests already in use (*a*) in a confirmatory sense, or (*b*) it may measure something not already taken into account by tests currently used. It may be useful in either case. Intrinsically, a test must show, by the results that it gives,—(*c*) that it actually does give a measure of something worth knowing, and (*d*) it must give a fairly reliable result. This last (*d*) has been accomplished, if measurements on the same individual, under conditions as nearly the same as practical considerations warrant, will, at different times, give not too widely divergent results. The condition (*c*) is met (granting that we aim to test individuals) if results from different individuals prove to

be different by variations of a larger order of magnitude than the variations in the same individual under identical conditions. If we aim to detect changes in the sensitivity of the individual, the condition (*c*) will have to be met somewhat differently.

The first point to investigate is the reproducibility of the result. It was shown in the previous communication upon this method, that for the three subjects, each having had some pactice as such, the variability of the result of a single series, representing about fifteen minutes' expenditure of time, was represented by a probable error of one unit or less on the arbitrary scale, equivalent to not more than about 8 or 9 per cent. in actual time.[1] The figures are given here in Table I.

This would be considered good reproducibility for such work as the present. But it must be remembered that these subjects were all more or less practiced in the method before the above results were obtained.

Before considering the results of the group, it may be well to examine the measurements taken upon the above three subjects in another respect. The question arises in connection with a test of this sort, not only what will happen when a measurement is attempted upon a new and unpracticed subject, but also what happens in the case of any subject, practiced or not, when several series are run in

[1] The relation between T (stimulus number) and t, the time of exposure (thousandths of a second) is given by the formula: $t = 200(\frac{1}{2})^{T/8}$. An extract from the original table embodying this relation is given here:

T	Time	T	Time	T	Time	T	Time
20	35.4	22	29.7	24	25.0	26	21.0
21	32.4	23	27.3	25	22.9	27	19.3

From these can be obtained the time values corresponding to the other numbers, if we remember that an increase of 8 units in T means that the time is divided by 2 and *vice versa*.

Furthermore, owing to the fact that the screen opening (area, 14.421 sq. mm.) was somewhat different from its designed size ($10\sqrt{2}$ sq. mm.) a slight correction is necessary to reduce the values for the group of subjects treated in this paper to standard conditions. This correction (to be perhaps unnecessarily accurate) consists in *decreasing* T by 0.22; or *increasing* the time by 2.0 per cent.

succession. There is the possibility of a fatigue effect, or of an effect of learning or practice. These two may be expected to affect the result in opposite ways, and may predominate, one or the other, irregularly; or may possibly tend to offset each other.

The average results for the three subjects first discussed follow, accompanied by the number of items entering into the average and by the mean variation. The average values

TABLE I

MEASURES OF FLUCTUATION OF THE RESULT OF A SINGLE SERIES FOR THREE PRACTICED SUBJECTS

T—Threshold, on arbitrary scale, reduced to standard aperture (area = $10\sqrt{2}$ sq. mm.).

$\sigma_{T/M}$—Standard deviation, relative to the mean for like conditions.

E_T—Probable error, in terms of arbitrary units, as T.

The last three columns give the times corresponding to T, and $T \pm E_T$ respectively.

Subject	T	$\sigma_{T/M}$	E	Time, Thousandths	
				Mean	Limits of *P.E.*
B.....	23.70	0.062	.99	25.7	23.6 to 27.9
C.....	26.85	0.056	1.01	19.5	17.9 to 21.3
L.....	23.48	0.049	.78	26.2	24.5 to 28.0

(M) are all in the neighborhood of unity, because, for the purpose of comparison, each item was, before averaging, divided by the mean of all results (T) for identical physical conditions. The figures given, both threshold and mean variation, are therefore expressed in terms of the general mean. Similar results, for a group of 33 unpracticed subjects (the first of a group of 96 subsequently to be discussed), appear in the fourth line. Here each item has been reduced to the average of the individual's four consecutive series as unity. (Table II.)

It will appear from the above that on the whole the variabilities of the successive results diminish somewhat, although in the case of the unpracticed subjects the change is in the opposite direction, but not to a significant extent. It appeared, however, that it would be out of the question,

in using the method as a test, to expend the time necessary to run four or even three series as a routine matter. That consideration will therefore be dropped, and only the first and second series will be considered in what follows.

The thirty-three subjects just mentioned were the first of a series of 101, which will form the basis of the major part of this discussion. Visual acuity was taken, at the time

TABLE II

COMPARATIVE VALUES AND VARIABILITIES OF THE RESULTS OF FOUR CONSECUTIVE SERIES

B, *C*, and *L*, are three practiced subjects; and the last row gives the results for 33 unpracticed subjects.

N—Number of series concerned in the average.

M—Mean of all values, relative to the mean of all series for the same subject.

MV—Mean variation expressed in the same terms as *M*. These values, multiplied by 1.253, should be comparable with the values of σ stated in the third column of Table I. The four series consumed an hour's time, more or less, according to the practice and natural aptitude of the subject.

Subject	Series 1			Series 2			Series 3			Series 4		
	N	*M*	*MV*	*N*	*M*	*MV*	*N*	*M*	*MV*	*N*	*M*	*MV*
B.....	18	.996	.048	18	1.002	.069	18	1.003	.049	18	.993	.037
C.....	22	.985	.057	22	.998	.058	22	1.005	.042	22	1.012	.046
L.....	21	1.016	.058	21	1.020	.036	21	.981	.042	21	.964	.030
Various..	33	.993	.044	33	.992	.046	33	1.019	.054	33	.996	.050

of the retinal sensitivity measurement, in 100 of these with each eye separately, and in 91 of these, with both eyes together. Unfortunately, in 5 cases of the 91, retinal sensitivity proved to be too low to measure with the apparatus as constructed, and rather than deviate from the established technique, it was decided to record these simply as 'too low to measure.' Some idea of the probabilities in these five cases may be gained from the fact that the lowest measurement actually made in conformity with the established technique was $T = 12.7$, and the three next 14.3, 14.5, and 14.7, respectively.

The distribution of 192 measurements, two upon each of

96 subjects, is shown in Fig. 1. The small circles represent values of the 'sliding' averages of four of the original class-

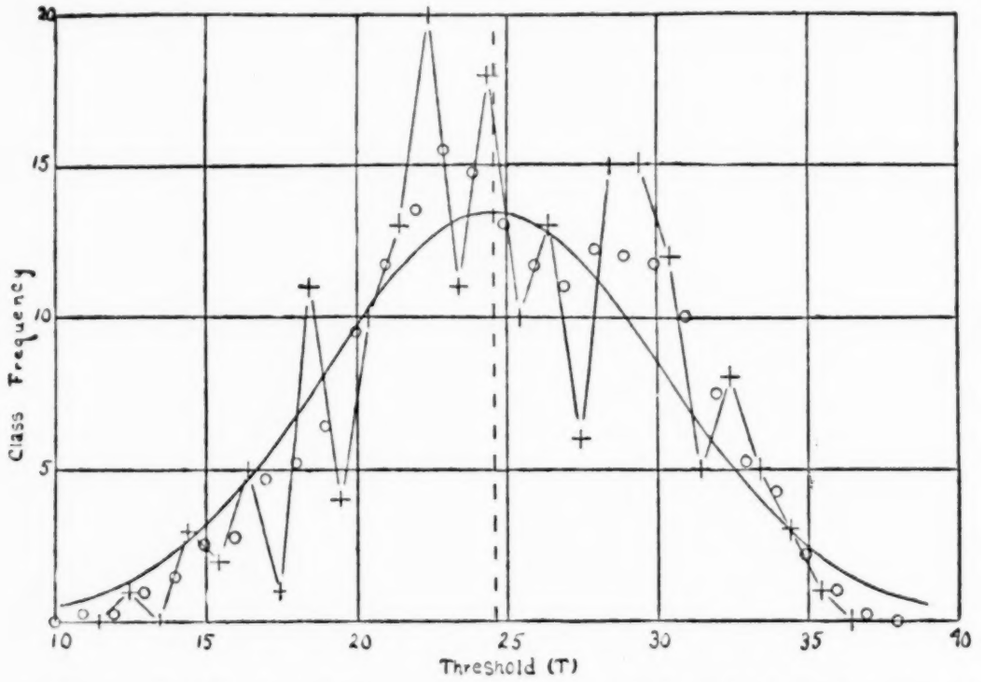

FIG. 1. Distribution of 192 measurements of retinal sensitivity; the values obtained from the first and second consecutive series of each of 96 different subjects. Abscissæ are (*T*). Class interval = 1. Class types are integral values + 0.4. Ordinates are class frequencies.

(*a*) Crosses and solid line—Frequency by classes.

(*b*) Circles—'Sliding' averages of four class frequencies.

(*c*) Smooth curve—Closest fitting normal distribution.

Ten of a possible 202 measurements fail to appear in (*a*) and (*b*) owing to the fact that they were too low to be measured with the apparatus used. These were taken into account in computing the constants of the normal distribution (*c*). See text.

frequencies and the smooth curve is that of a normal distribution based upon characteristics of distribution derived as follows: (1) A mean drawn from the whole of the 192 measurements less the 10 highest, omitted to offset the absence of the corresponding 10 lowest of 202 possible measurements, which are not included in the 192 because of their being too low to measure; and (2) a measure of dispersion derived from the sum of the deviations of that fraction of the distribution lying above the mean.

The first of the curves in the diagram (*a*) exhibits irregularities due to the method of interpolation by which a numerical threshold was derived from the results of one series. This method favored the even integers; that is, of the 192 inter-

polated values, 119 were even integers plus a fraction, and 73 only were odd integers plus a fraction. Inspection of the figure will show this irregularity. However, the circles representing the average frequency of four consecutive classes, plotted against the average of the corresponding four class-types, represent values from which this systematic irregularity is eliminated, and among which other irregularities are also, in part, evened out. Nevertheless, in comparing the line of the circles with the curve of normal distribution, it would appear that there are in the former two maxima: one at about the point where $T = 23$, $(t = 27\sigma)$ another at $T = 29$, $(t = 16\sigma)$; with a minimum between them at $T = 27$, $(t = 19\sigma)$. In other words, the plot suggests very strongly that the distribution of results is a bimodal one.

The relation existing between the results of the first and second series on each of the 96 subjects is shown in Fig. 2,

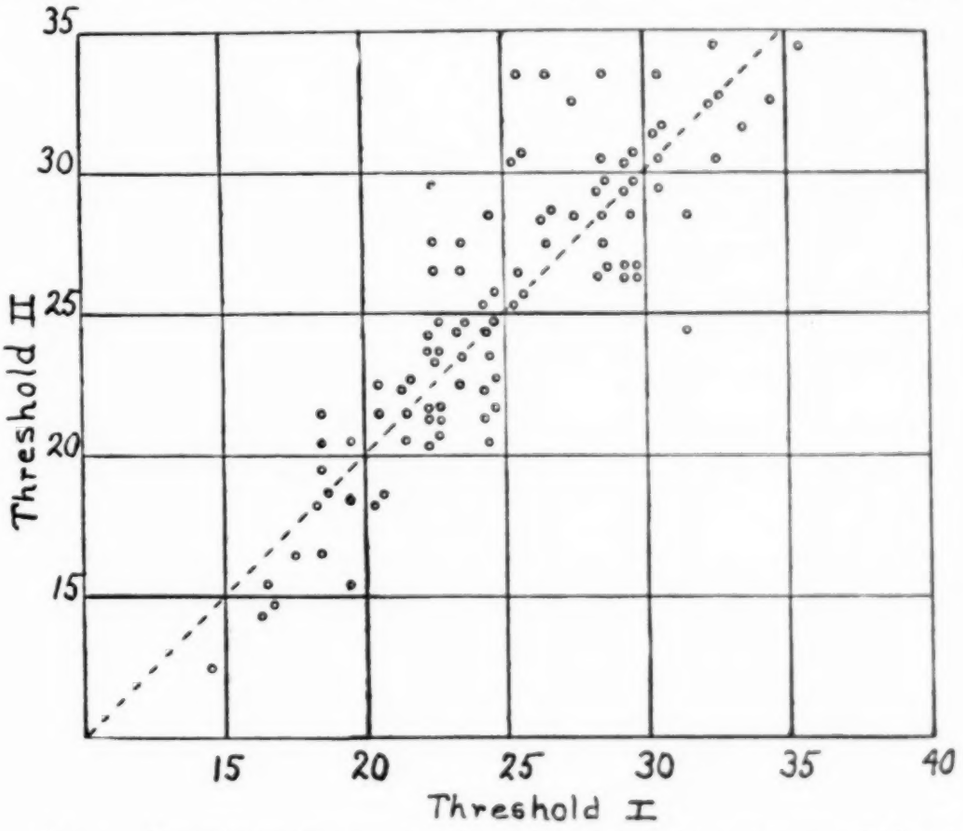

FIG. 2. The relation between the results of the first (I.) and second (II.) series, for each of 96 subjects. The oblique dotted line represents the condition of equality. Those plotted below this line gave a result on the second series lower than that of the first, and *vice versa*.

where the second of these is plotted directly against the first. The two means are I, 24.91 and II., 25.17; very close together. The oblique dotted line, drawn at 45° inclination, is the locus of the cases in which the two are exactly equal. The mean values just stated would indicate a tendency in the subjects to go higher on the second series, and a glance at the figure will show this tendency in general to be greater in those above the mean than in those below it and, in fact, to be reversed in a few cases at the other extreme. This shows itself also in the two standard deviations. That for the first series is 4.6, for the second 5.1. The correlation-ratio is 0.857 ± 0.018.

The characteristic equations derived from the above are:

$$x_1 = 0.77x_2, \qquad E_1 = 1.6; \text{ and}$$
$$x_2 = 0.96x_1, \qquad E_2 = 1.8;$$

in which the two probable errors (E) may be taken as an index of the variability of these inexperienced subjects, as a group, for the first and second consecutive trials respectively. This indicates a lower degree of reproducibility for the group of 96 subjects than for the three considered in Table I. There the values for E are of the order of 1.00 or less. The group of 33 various observers will be seen to compare very favorably with the three observers, *B*, *C*, and L (Table II., *M. V.*' s).[1] They therefore appear to show a better reproducibility than the group of 96 as a whole, of which they are a part.

This is incidental to the rather unusual way in which the method of correlation [2] has here been applied, for the purpose of estimating, from the results of a group, the expected standard deviation of repeated identical measurements upon an individual. In plotting the second set of results, *b*, against the first (*a*) as abscissæ, it will be seen that although the *a*'s are of identical class type for any column of *b*'s, yet in the case of any pair plotted in the column, the *a* is still in every case a variate from a possible mean, from which it deviates, on the whole, by an amount represented by its

[1] The probable error for all results of the group of 33, Table II., is 1.05, on the same basis as the values for E in Table I., for subjects *B*, *C*, and *L*.

[2] The correlations discussed here were performed by Pearson's method.

own standard deviation. Thus the standard deviation from their characteristic line of the b's ($\sigma_{2.1} = \sigma_2\sqrt{1 - r_{12}^2}$) is not the standard deviation to be expected of repeated identical measurements, but is greater than the latter, as the standard deviation of a (sum or) difference exceeds that of either term; by the factor $\sqrt{2}$ when the two are of equal variability. Accordingly the two E's stated above should be divided by $\sqrt{2}$, and they then become $E_1 = 1.1$, $E_2 = 1.3$, which conform fairly well to the results stated in Tables I. and II.

Of especial interest in determining the value of this test is the question of the correlation existing between its results and those of the ophthalmologic tests. And of these, the relation with visual acuity, taken by the Snellen letters, was of especial interest. Visual acuity was taken for each eye separately, and for binocular vision, by hanging a test-letter card upon the screen of the apparatus, so that the conditions of the retinal sensitivity test were exactly reproduced for the letter test. The results of the correlations, performed by Pearson's method, are given in Table III. An incidental question, worked out by way of side issue, is the relation of binocular visual acuity (V_0) to the two values of monocular visual acuity obtained with the better eye, (V_1), and the poorer eye, (V_2), respectively. The equation for this value, derived from the correlation-ratios of the first order and the standard deviations of the second order, involving the variables 2, 3, and 4 only, is as follows:

$$V_0 = 0.75V_1 + 0.16V_2 + 0.18, \qquad (1)$$

the standard deviation of V_0 is $\sigma_{2.34} = 0.13$. It is to be remembered that the bulk of the results fell about the region of $V = 20/30$ to $V = 20/15$ (6/9 to 6/4.5 metric or 1.00 to 1.33 decimal) and the application of this equation is not safely to be extended outside these limits.

The interpretation of a characteristic equation such as the above is that V_0, calculated from the observed values of V_1 and V_2, is the expected value of binocular acuity, and that experimentally determined values may be expected to deviate from this as implied in the accompanying standard deviation, $\sigma_{2.34} = 0.13$.

Similarly, the characteristic relation between retinal sensitivity (T) and the three values of visual acuity above enumerated is:

$$T = 5.8V_0 - 2.3V_1 + 4.0V_2 + 16.2; \qquad (2)$$

and the standard deviation of T is $\sigma_{1.234} = 4.24$; the same re-

TABLE III

THE RESULTS OF CORRELATION OF RETINAL SENSITIVITY (1) WITH VISUAL ACUITY (2, 3, AND 4) MEASURED WITH THE SNELLEN LETTERS UNDER THE CONDITIONS OF THE EXPERIMENT

The results from the letter chart have been reduced to decimal values (*e.g.*, 20/20 = 1.00; 20/15 = 1.33; 20/20 + 5/8 = 1.00 + 5/8 × 0.33 = 1.21) and the values of σ in division *c* of the table are expressed in such decimal units. Retinal sensitivity is expressed in arbitrary scale units, the same as those in which T is elsewhere expressed.[1]

ss—subscript indicating the variables involved and their interrelation.

N—number of cases involved.

r—correlation-ratio.

σ—standard deviation.

(*a*) *Numerical Designations of the Variables:*

1. Retinal sensitivity (mean of I. and II. series.)
2. Visual acuity (Snellen), binocular vision.
3. Same, better eye alone.
4. Same, poorer eye alone.

(*b*)

Correlation-ratios

ss	*N*	*r*, Zero Order	E_r	*ss*	*r*, First Order	*ss*	*r*, Second Order
12	86	0.375	0.062			12.34	0.177
13	95	0.321	0.062			13.24	−0.068
14	95	0.355	0.060			14.23	0.147
23	91	0.871	0.017	23.4	0.716	23.14	0.716
24	91	0.723	0.034	24.3	0.191	24.13	0.160
34	100	0.760	0.029	34.2	0.384	34.12	0.389

(*c*)

Standard Deviations

ss	*N*	σ, Zero Order	*ss*	σ, Second Order	*ss*	σ, Third Order
1	95	4.62			1.234	4.24
2	91	.271	2.34	.131	2.134	.129
3	100	.275	3.24	.125	3.124	.125
4	100	.247	4.23	.157	4.123	.155

[1] For a more complete explanation of the meaning of these symbols and of the methods of correlation see: Yule, An Introduction to the Theory of Statistics, Fifth Edition, London, 1919; Chap. IX ff. and Chap. XII.

servations as to the limits of applicability of this equation to apply, as in the case of equation (1).

These relations exhibit one or two points worthy of note. From equation (1), it will be noted that binocular visual acuity depends in a major degree upon the vision of the better eye. Whether it depends at all upon the vision of the inferior eye is an important question. V_2 being ordinarily somewhere about unity, the value of the V_2 term will be about 0.16, which is not greatly in excess of the standard deviation of V_0 from its value computed by the equation. However, while the plotted results show many cases where binocular vision was distinctly better than the vision of the better eye alone, there are but 4 cases in a total of 91 where the reverse is true, and this only to the extent of one class interval, equal to two single letters on the chart, or decimally expressed, equal to a difference of from 0.04 to 0.15 according to the portion of the chart involved. The mean value for V_0 is 1.22; for V_1, 1.16; and for V_2, 1.06; which also would indicate a definite extent to which, on the average, vision of the one eye is helped by the second, inferior, eye in binocular vision.

As to equation (2), the negative coefficient of V_1 would seem to indicate that the better eye has a negative contributory value to retinal sensitivity. By the same reasoning as the foregoing we see that in general the amount of the second term, V_1 being about unity, does not exceed the standard deviation of T (4.24). We must also bear in mind that the three V's bear high correlations, each to the other. The writer would interpret the equation as meaning that the better eye contributed to the result in the case of retinal sensitivity (that is, reinforced the poorer eye) only in so far as it does the same with reference to the letter chart; that is, only in so far as it coöperates with its inferior fellow in other respects. Retinal sensitivity is shown by the equations to be more highly dependent upon V_0 than upon either V_1 or V_2; which is perhaps to be expected, for retinal sensitivity was also taken binocularly.

This equation answers the question as to whether this test is to be looked upon (as indicated in an early paragraph

of this paper) as supplementary to other tests in (*a*) a confirmatory sense or (*b*) as measuring something not already taken into account.

For the measures of retinal sensitivity in the 95 cases taken into account in the correlations, the standard deviation (σ_1) is 4.64. Equation (2) should give for T the closest value possible to compute from all the data on visual acuity. The variations of the actual measurements from this computed value are represented by $\sigma_{1.234}$, following equation (2), and equal to 4.24. Thus the attempt to express retinal sensitivity in terms of visual acuity is futile. It still deviates in much the same degree from the computed values, and we must conclude that retinal sensitivity, measured by the present method, involves a new characteristic of vision that is not to any significant extent taken into account in the measurement of visual acuity by means of the letter-chart.

In examining the relation of the retinal sensitivity results to the ophthalmologic findings as a whole, it was necessary to establish some reasonable means of numerical rating as to the latter. The ophthalmologic examination, of which the results were used, was part of the general medical examination of candidates for flying status, known at the Medical Research Laboratory as '609,' from the form number of the blank used to record the results. The criteria of ophthalmologic qualification, as practiced at the time this work was done, are here reproduced. The method of rating, as will be seen, was to score 1 against the candidate for a contingent disqualification, as implied in the following paragraphs, 3, 5, 8, and 9; to score 2 for disqualification in one particular; and to score 1 for each further point of disqualification. The rules following were interpreted literally. Following each paragraph, in parenthesis, are such supplementations and changes as were necessary to make the rating just and unequivocal, or remarks necessary to explain the nature of the test to one unfamiliar with the special practice.

1. *Visual Acuity.*—The minimal visional requirement for each eye is 20/20. If two or three letters are not read in the 20/20 line they may be offset by an equal number of letters read in the 20/15 line.

(The ratings taken in the experimental room were used, not those furnished by the ophthalmologist. In view of the qualification in the foregoing paragraph, a subject was passed if he read 20/20 less two letters, assuming in the absence of record of the fact, that he was able to offset these by two read in the 20/15 line. The decimal value 0.92 or over satisfied this condition.)

2. *Depth Perception at 6 Meters.*—An average depth difference of more than 30 mm. disqualifies the applicant. (Two rods, 1 cm. in diameter, separated 6 cm. laterally, must not be more than 30 mm. apart in the direction of vision when set to apparent equidistance by the candidate. The rods were seen in silhouette against a bright background, and seen through an opening which did not permit vision of either top or bottom of the rods.)

3. *Maddox Rod Screen Test at 6 Meters.*—Esophoria of more than 4^{Δ} is a disqualifying factor if associated with less than 4^{Δ} of prism divergence, or if associated with diplopia in the lateral positions on the tangent curtain, or if associated with an amount of accommodation near the lower limits, or if associated with an amount of hyperopia near the disqualifying limit.

4. Esophoria of more than 10^{Δ} is a disqualifying factor, even if unassociated with any of the preceding conditions.

5. Exophoria of more than 2^{Δ} is a disqualifying factor if associated with an angle of convergence near the disqualifying limit, or if associated with diplopia in the lateral positions on the tangent curtain.

6. Exophoria of more than 5^{Δ} is a disqualifying factor even if unassociated with any of the preceding conditions.

7. Hyperphoria of more than $\frac{1}{2}^{\Delta}$ disqualifies the applicant without further supporting evidence.

8. *Maddox Rod Screen Test at 33 cm.*—Exophoria of 4^{Δ} may be considered the normal condition. Any considerable variation from this condition is to be interpreted in connection with the other associated tests. An exophoria of more than 12^{Δ} at 33 cm. disqualifies.

(1 to 7^{Δ} exophoria at 33 cm. were taken as the normal condition. Outside of these limits a question was scored.)

9. *Prism Divergence.*—Prism divergence of more than 9^{Δ} disqualifies the applicant if associated with an angle

of convergence near the disqualifying limit. If less than 4^{Δ} of prism divergence is found associated with more than 4^{Δ} of esophoria at 6 meters, the applicant is disqualified.

(Those grades of muscle abnormality which do not disqualify except contingently upon some other finding were marked with a question. One or more questions in the absence of final disqualification scored 1 only against the applicant.)

10. Prism divergence of more than 15^{Δ} disqualifies without further supporting evidence. Prism divergence of less than 2^{Δ} disqualifies without further evidence.

11. *Test of Associated Parallel Movements.*—The applicant is disqualified if the underaction or overaction of any of the extrinsic ocular muscles causes diplopia except in the extreme positions, where a small separation of the images may be disregarded. Nystagmus disqualifies if it is demonstrated except in extreme positions.

12. *Inspection of the Eyes.*—Any pathological condition which may become worse or interfere with the proper functioning of the eyes under the fatigue and exposure of flying, disqualifies the applicant.

13. *Accommodation.*—Accommodation is normal if it lies between limits 2 diopters above and below the mean for the applicant's age. Failure to read within these limits disqualifies.

TABLE OF MEAN VALUES OF ACCOMMODATION POWER

Age	Dptrs.	Age	Dptrs.	Age	Dptrs.	Age	Dptrs.
18	11.9	25	10.2	31	8.6	37	6.8
19	11.7	26	9.9	32	8.3	38	6.5
20	11.5	27	9.6	33	8.0	39	6.2
21	11.2	28	9.4	34	7.7	40	5.9
22	10.9	29	9.2	35	7.3	45	3.7
23	10.6	30	8.9	36	7.1	50	2.0
24	10.4						

14. *Angle of Convergence.*—An angle of convergence smaller than 40° disqualifies,

15. *Central Color Vision.*—If it is apparent that mistakes made by the applicant are due to color confusion and not to carelessness or failure to understand instructions, he is disqualified.

(Color-vision was left out of consideration in grading the subjects. There were but three cases not definitely found normal as to central color vision as follows: (1) no record

as to color vision. Record otherwise incomplete, and subject therefore excluded from consideration; (2) no record as to color vision, disqualified on account of visual acuity and hyperphoria; (3) found to confuse reds and greens. No other disqualification or question.)

16. *Field of Vision for Form and Color.*—The normal visual field for form is largest; those for blue, red, and green are successively smaller in the order given. The color fields should be nearly concentric with the form field. Any marked contraction of the form field disqualifies the applicant for flying.

17. *Refraction.*—The applicant is disqualified if he cannot read 20/20 without more than one diopter of correction, either hyperopic, myopic or astigmatic.

Concerning paragraphs 11, 12, and 16, it was assumed that no disqualification existed unless specified. The records furnished the writer did not ordinarily state the findings on these points.

Of a total of 101 subjects upon whom retinal sensitivity measurements were made, there were 59 with records complete as indicated above, and only these are included in what follows.

It will be seen from the plot (Fig. 3*a*), that those showing the highest number of points as to ophthalmologic defects have, in general a low retinal sensitivity; curve *b* shows the average number of points of defect for each class as to retinal sensitivity; and curve *d* the average retinal sensitivity for each particular score. Both of these curves bear out the first statement made. The correlation ratio is -0.444 with a probable error of ± 0.070.

If, however, we exclude those having an ophthalmologic score of 4 or more against them, we get a correlation ratio of 0.009 ± 0.094 which is no correlation at all. The curve (*c*, Fig. 3) shows that there is no significant relation between retinal sensitivity and the ophthalmologic findings when those defective with a score of 4 or more are excluded.

We must conclude, therefore, that there is no relation between the defects elicited by the ordinary ophthalmologic examination on the one hand, and the present retinal sensi-

tivity rating on the other; excepting those cases in which the clinical defects are comparatively numerous, in which these are signs of some rather general, perhaps distinctly pathological, derangement of the visual organs which comes to be generally reflected in their various functions, of which retinal sensitivity is one.

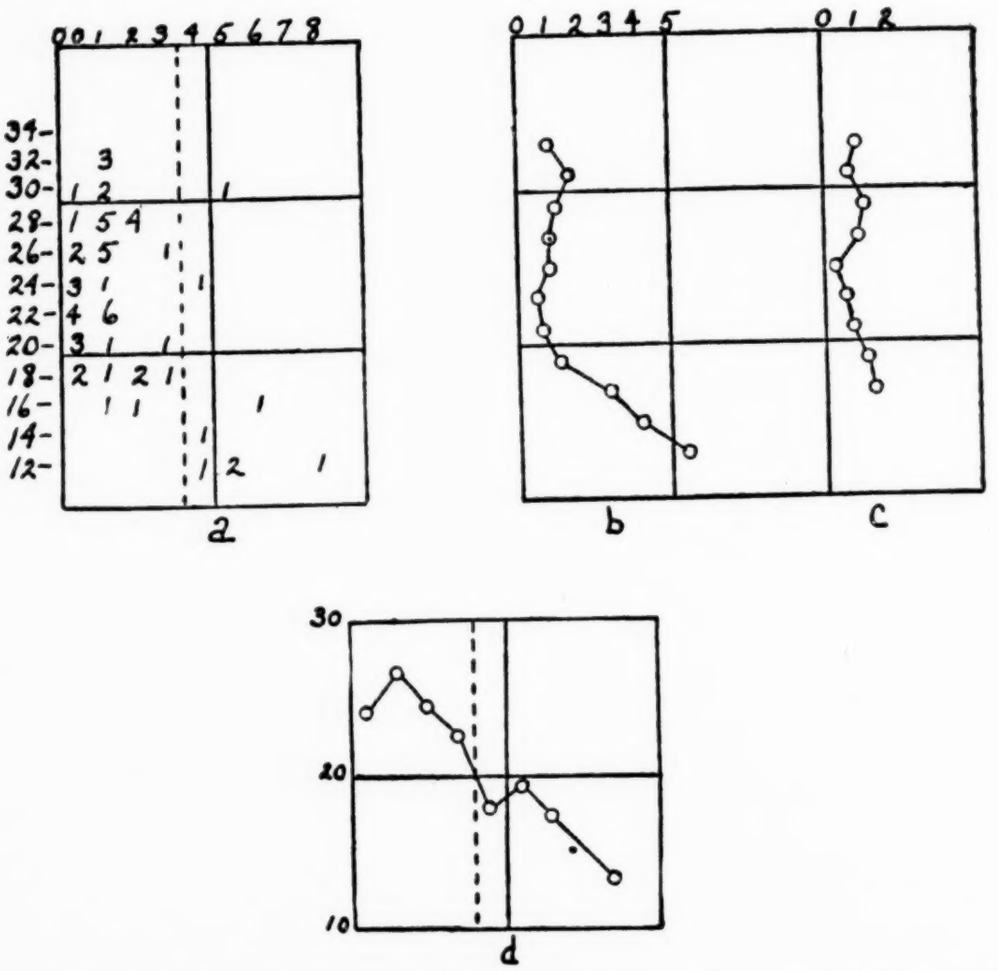

FIG. 3. The relation between retinal sensitivity measurements and the general degree of visual defectiveness found in the ophthalmologic examination.

(*a*) Classified results. Ordinates indicate class as to retinal sensitivity. The class-interval is 2 units on the scale, and the mean of the class is greater than the designation given by one unit. *E.g.*, in the 30-class, the lowest possibly included is 30.00, the highest 31.99, the mean (or type of the class) therefore is 31.00 minus. Abscissæ, points disqualification; see text.

(*b*) Average points of disqualification for each retinal sensitivity class.

(*c*) Same as (*b*), scores of 4 or over excluded.

(*d*) Average retinal sensitivity (ordinate) for each class as to points disqualification.

The cases falling to the right of the dotted line in (*a*) and (*d*) are those excluded from (*c*).

This conclusion will perhaps be borne out by inspection of the following synopsis of all results in the 59 cases considered. Where retinal sensitivity was too low to measure this is indicated by the word 'low' in place of the numerical result. Otherwise the synopsis should be self-explanatory:

Summary of Cases

Serial No.	Ret. Sens.	Clinical Description		No. of Cases
—	—	*Unquestioned, Score 0*		16
—	—	*Questioned, Score 1*		
		Esophoria	5	
		Exophoria	22	
		High prism divergence	7	
			34	
		Counted twice or more, less	9	
			25	25
		Just disqualified, Score 2		
101	29.8	Visual acuity	1	
62	17.0	Depth perception	1	
80	28.2	Exophoria at 6 meters	1	
19	19.9	Low convergence	1	
90	28.6	Hyperopia, 1 diopter, esophoria	1	
100	28.7	" 3 "	1	
16	18.6	Hyperopic astigmatism, corrected	1	
			7	7
		Disqualified, Score 3		
42	19.0	Visual acuity and hyperphoria	1	
83	21.7	Hyperopia, low accommodation	1	
91	27.8	Hyperopic astigmatism, low accommodation	1	
			3	3
		Disqualified, Score 4		
		Hyperopic astigmatism, poor depth perception, with:		
55	15.6	Convergent strabismus	1	
46	Low	Low visual acuity	1	
82	25.9	Deficient accommodation	1	
			3	3
		Disqualified, Score 5		
87	31.1	Hyperopia, exophoria, hyperphoria, low convergence and accommodation	1	
85	Low	Hyperopic astigmatism, low visual acuity and depth perception, hyperphoria	1	
77	Low	Hyperopic astigmatism, low visual acuity, depth perception and accommodation	1	
			3	3

Serial No.	Ret. Sens.	Clinical Description	No. of Cases
		Disqualified, Score 6	
39	17.5	Myopia, corrected; low visual acuity, depth perception and convergence; exophoria..... 1.....	1
		Disqualified, Score 8	
63	Low	Myopic astigmatism, strabismus, low visual acuity, accommodation and depth perception. Impossible to measure convergence..... 1.....	1
		Total number of cases.....	59

Summary of Conclusions

1. Measurements of retinal sensitivity made upon 101 subjects by a new method, previously described in detail, show individual differences far in excess of the accidental variations for the same individual and the same conditions.

2. The result of any one of as many as four consecutive series is of about equal reliability even with unpracticed subjects.

3. The distribution of the various results of retinal sensitivity measurement suggests a natural division of individuals into two groups, most numerous where the times of stimulation are approximately 27 and 16σ respectively.

4. Correlation of the results of the first series with those of the second, taken immediately after the first, in 96 subjects, shows almost identical characteristics for the two. The correlation ratio is 0.857 ± 0.018.

5. The characteristic equation derived from correlation of the retinal sensitivity results (average of first and second series) with visual acuity indicate that retinal sensitivity is somewhat dependent upon the visual acuity of the eye with poorer vision, but not upon that of the better eye unless the advantage of this latter is evident in enhanced binocular visual acuity.

6. Due allowance for the differences due to different visual acuity leaves individual differences in retinal sensitivity measurements unaccounted for, almost to the original amount. The standard deviation from 4.64, is reduced only to 4.24 when correction is made for the differences due to different visual acuity.

7. Correlation between retinal sensitivity and a rating based upon the ophthalmologic portion of the Air Service medical examination for flying status, indicates that in general no definite relation exists between retinal sensitivity and disqualification; except when the latter is based on findings which, in degree and character, indicate distinctly pathological conditions.

8. Incidentally, the method of correlation indicates that binocular visual acuity is dependent chiefly upon the eye which, used singly, shows the better vision, to the extent of nearly five times its dependence upon the eye with the poorer vision. This has exclusive reference to results obtained with the letter-chart.

Appendix

The apparatus used in this work differed from that used in the former work in the substitution of a perforated and painted metal slip for the punched card formerly used for the test-stimulus aperture. It was found that simply drilling and painting the slip would not suffice, since the paint tended to thin itself, by its surface tension, at the margins of the aperture, causing a ring-shaped shadow by which a photometric balance between screen and aperture was rendered impossible.

The following will give an idea of the technic found necessary in preparing replaceable apertures free from such defect:

Sheet zinc of 0.017 inch thickness was found suitable for the purpose. The hole at the center was drilled to admit the end of a reamer (known to the trade as a 'No. 1 taper pin reamer'), provided with a wooden sleeve adjusted so that the hole, reamed from the back until the sleeve just touches the metal, was of just the right size to fit a carefully sized gauge plug. This plug was made for the purpose and measured 0.1687 inch (4.285 mm.) in diameter. The margins of the aperture were then freed from burrs by passing a moderately fine file flat over the front and back surface of the metal.

Before painting the slips, the back of the hole was covered

with surgeon's adhesive, and the hole filled with Alabastine (water) paint, mixed to a putty consistency. Beeswax was tried and worked almost as well. The filling was pared down flush with the face of the zinc by means of a clean-edged carpenter's chisel, and the face of the zinc cleaned free of excess filling material at the same time.

The face of the slip was then painted. Two coats of a good ordinary white oil paint were applied in the usual way, and as these nearly covered the metal, a coat of alabastine white was applied. The filling was then removed by first stripping the adhesive from the back and locating the center of the hole by pricking with a needle, and then carefully punching the filling out, from the face, piecemeal, without touching the parts which would form the margins of the hole. After this the reamer was run into the hole from the back, as in the first place, carefully and *only* until the collar on the reamer lightly touched the zinc.

The object of this procedure was that the edge of the opening should be perfectly sharp, and that the surface of the face of the slip should continue to be plane, accurately and exactly, until it met the cylindrical (slightly conical) surface cut by the reamer. If carefully and successfully done, it was readily possible to make the aperture disappear by adjusting the brightness of the back screen.

Considerable care had to be used in handling the finished slips, as they were rendered useless by the slightest soiling or chipping of the paint at the margins of the aperture.

ADULT TESTS OF THE STANFORD REVISION APPLIED TO UNIVERSITY FACULTY MEMBERS

BY HELEN HUBBERT CALDWELL

Dayton, Ohio, May 8, 1922

Object of Research.—The present investigation was the outgrowth of one previously made on the application of the Stanford Adult Tests to College Students,[1] and was carried on in the spring of 1919 while the writer was instructor in educational psychology at the University of Wisconsin. Its object was threefold. First, to determine the possible limits of performance for adults especially chosen because of reputed brilliancy or genius. Second, by recording the percentage of failures in each part of the test and relating them to the performance as a whole, to approach a better evaluation of the separate tests. Third, to test the validity of the assertion that there are significant differences in the results obtained from persons inclined to literary pursuits and persons inclined to scientific pursuits, the former being markedly superior in the so-called Verbal Tests, while the latter excel in what are called the Constructional Tests.

Securing Subjects.—The writer had not realized until she attempted to secure subjects for this investigation, how widespread and ofttimes violent among cultured and educated persons was the opposition to mental testing of any sort. Attempts to get at the reasons for such opposition were in most instances fruitless. Beyond the stock objections, "mind cannot be measured," and "mental tests take no account of individuality," not many would express themselves. They were content to condemn mental testing on general principles and refuse to become subjects 'even in the interests of science.'

[1] Caldwell, Helen Hubbert, 'Adult Tests of the Stanford Revision Applied to College Students,' *Jour. Ed. Psychol.*, Dec., 1919, Vol. X., No. 10, p. 477.

The first intention, which was to have only heads of departments and full, associate and assistant professors as subjects, had to be abandoned, and in four instances it was necessary to go below the rank of instructor. Records for three student assistants and one student fellow are included, but each of these students had an unusually brilliant record in his particular department, and might be classed as a 'genius' in the ordinary meaning of that term. Neither was it possible to secure as many subjects as desired. The original plan of the research called for thirty subjects in each group, sixty in all. Slightly over half of that number were actually obtained.

So far as was possible, faculty members chosen were those whose predilections for scientific or for literary pursuits were well established. The effectiveness of personal appeal as against telephonic communication was demonstrated in the securing of subjects. Attempts to arrange for the tests by telephone were without exception unsuccessful. Personal interviews with about seventy faculty members, in which the purpose of the investigation was explained, resulted in but thirty-six subjects.

Opinions Regarding the Tests.—Reasons—or shall we say excuses—for not coöperating were plentiful and some of them quite amusing. Most frequent was the plea, "I haven't time," and this was allowed to stand even after it was explained that but 45 or 50 minutes would be required. In nearly every case where coöperation was secured it was the declaration, "We aren't really testing *you*, you see; it's the *tests* we are testing; we want to see if they really do show differences between you literary folks and the scientists" (or *vice versa*, as the case might be), which finally won consent. Once the experimenter was soundly berated for even venturing to request the assistance of a very noted professor. She was told that his time was precious, that he was protected by University rulings from the annoyance of book agents, solicitors and peddlers and should be so protected from her intrusion. In another instance, refusal to take the tests was apparently based upon the pretense that it would

be such a 'waste of time,' since 'no mental tests were worth anything anyhow." But the same gentleman added, "I may be a fool, but I don't intend to let anyone find it out if I can help it!" No doubt in many cases refusal was due to an unacknowledged fear of failure and the functioning of a 'defense mechanism.'

By no means all of those who finally agreed to submit to the experiment were friendly to mental testing. In at least four cases the tests were declared to be 'utter nonsense.' One subject considered all mental tests 'vicious,' while another characterized them as 'piffle.' Attempts to draw these reagents out and get reasons for their opinions failed. In four instances the attitude was one of complete boredom, while three reagents submitted with an air of amused toleration. Three subjects were apparently willing enough, but seemed unable to concentrate on the work in hand. Six were willing, but extremely self-conscious. The remaining subjects were alert, willing, and coöperated exceedingly well, with one exception where there was utter failure to coöperate —fortunately the only case of its sort which was encountered. This record is not included in the averages or results, since only the average adult tests were attempted. Tests 1, 2 and 3 were passed, 4 was not passed, 5 was a complete failure, 6 the subject positively refused to attempt, saying, "I am not interested, don't want to put my attention on it." In alternate 1 '*a*' was missed but '*b*' was correct. '*a*' of alternate 2 was missed, '*b*' was correct, '*c*' was not attempted. Instead the experimenter was told, "If you want my real reaction to these problems they bore me stiff." At this point the test was discontinued!

One subject declared himself as unalterably opposed to all attempts at mental testing "because they take no account of genius, and do not enable one to get at the *urge* behind the actual performance." When told that psychologists themselves were working on this very problem in an attempt to get at this element of effort or 'urge' by devising character tests, he replied, "It can't be done. I myself have spent hours trying to devise something of the sort." Though con-

vinced that he could not 'get by' the immigration authorities at Ellis Island if required to take their tests, he emerged from the ordeal with the highest Intelligence Quotient it is possible for an adult to make, 121.8, and was the only one of the 36 subjects who did not fail on any part of the test. This instance and several others bear out Miss Downey's [1] statement that the reagent's score bore little relation to his own estimate of his ability.

The digit-span and other immediate memory tests called forth more objections than any others. In a number of cases, however, after the first refusal to try had been overcome, the tests were passed correctly. Several suggestions regarding the 'number tests' were made. Two subjects were confident they could pass them if they were presented visually instead of orally. A set of numbers so presented at the close of the test gave no better results with one of these subjects than the auditory presentation which is customary. The other subject, however, passed all three of the digit-span tests when the numbers were presented visually although he had failed on all three with the auditory presentation. A mathematics professor explained that mathematicians would not be apt to pass the digit-span tests because of their 'habit of thinking in wholes and generalities.' He himself passed them creditably. Only two of the subjects of the scientific group really objected to the 'number' tests, while seven of the literary group had to be coaxed to try them.

Results for all Subjects.—Two of the 36 records were discarded before making up averages, one in the literary group because of lack of coöperation, one in the scientific group because of six interruptions which occurred during the test. All tables, averages and conclusions drawn from the experiment are based upon the remaining 34 records. In most cases percentages and averages are given in whole numbers. The same notation is used as in the writer's previous paper on Stanford tests, the tests in the sixteen-year group being referred to as I—1, I—2, I—3, I—4, I—5, I—6, Al. 1 and

[1] Downey, June E., 'The Standard Adult Intelligence Tests,' *Jour. of Delinq.*, 1917, 2, 144–155.

Al. 2, and the eighteen-year-old tests being referred to as II—1, II—2, II—3, II—4, II—5, and II—6. Intelligence quotients are referred to as I.Q's.

The I.Q.'s for the entire list of faculty members tested range from 104.7 to 121.8. Only 23.5 per cent. fall below the Superior Adult level, and none below the Average Adult level. Nine per cent. rank as Very Superior. The average I.Q. is 114, 47 per cent. showing a higher I.Q. than this average while 50 per cent. fall below it. Seventy-six and five tenths per cent. reach or exceed the Superior Adult level. Time required for the tests varied from 22 to 60 minutes, the average time required being 40 minutes.

TABLE I

SHOWING, (*a*) THE NUMBER AND PER CENT. OF FAILURES ON EACH TEST FOR THE ENTIRE GROUP; (*b*) THE NUMBER AND PER CENT. OF FAILURES FOR THOSE HAVING A SUPERIOR ADULT RECORD; (*c*) THE NUMBER OF FAILURES ON EACH TEST IN MISS DOWNEY'S GROUP OF WYOMING UNIVERSITY FACULTY MEMBERS

	(*a*) Entire Group Failures		(*b*) Superior Adult Failures		(*c*) Superior Adult Failures (Downey)
Test	No.	Per Cent.	No.	Per Cent.	No.
I—1....	0	0	0	0	0
I—2....	0	0	0	0	5
I—3....	2	6	1	3	1
I—4....	13	38	8	24	6
I—5....	17	50	8	24	4
I—6....	8	24	3	9	13
Al—1....	19	56	10	30	Not recorded
Al—2....	2	6	1	3	" "
II—1....	0	0	0	0	1
II—2....	13	38	7	20	10
II—3....	15	44	6	18	11
II—4....	5	16	3	9	10
II—5....	17	50	7	20	7
II—6....	3	9	1	3	8

Reference to Table I. (*a*) will show the percentage of failures on each test. Attention is called to the relatively large percentage of failures on tests involving immediate memory only, I—5, Al—1, II—3 and II—5. Even when the

records of those below the Superior Adult level are eliminated this is still true. (See Table I. (*b*).) Note also the weakness on tests involving the use of the imagination, I—4 and II—2.

For a comparison of the rating of that part of the University of Wyoming faculty tested by Miss Downey with that part of the University of Wisconsin faculty tested, see Table I., (*b*) and (*c*), where the number of failures recorded in each test for those with an I.Q. indicating superior intelligence is tabulated.

Comparison of these results with those secured in the work with college students is interesting. (Table II.) In tests largely involving the verbal element the faculty ratings are distinctly higher—(I—1, I—2, I—3), in immediate memory tests [1] they are as low or lower—(I—5, Al—1, II—3, II—5), in the two tests which depend almost entirely upon use of the imagination (I—4, and II—2), faculty and students rate about alike. On the code test the faculty is slightly better, (I—6) for the logical memory and ingenuity tests, (II—4, II—6), the faculty record is considerably better, while the physical relations test, (Al—2), shows almost as

TABLE II

SHOWING THE PERCENTAGE OF FAILURES ON EACH TEST FOR THE SOPHOMORES AND JUNIORS OF RANDOLPH-MACON WOMAN'S COLLEGE AND THE FACULTY GROUP TESTED AT THE UNIVERSITY OF WISCONSIN

Test	Percentage of Failures Recorded		
	Sophomores	Juniors	Faculty
I—1	33⅓	12½	0
I—2	12½	12½	0
I—3	25	12½	6
I—4	37½	37½	38
I—5	33⅓	29⅙	50
I—6	29⅙	29⅙	24
Al—1	58½	62½	56
Al—2	83⅓	62½	6
II—1	70	50	0
II—2	25	42	38
II—3	42	12½	44
II—4	46	33⅓	16
II—5	54	70	50
II—6	50	62	9

[1] Miss Downey's 'digit-span tests' + Al.1.

great a contrast between faculty and students as the verbal tests. Comparing these results with Miss Downey's we also find the faculty to excel on vocabulary. We cannot agree that they are better on digits reversed, nor does it appear that our student group tested excelled on code, paper-cutting, digits forward or logical memory. In fact for the code and logical memory tests quite the reverse is true. It would not be safe to draw general conclusions regarding reactions of students as compared with faculty from either set of results.

Results for Separate Groups.—We turn next to a consideration of results obtained from the two separate groups in this problem. There were 17 subjects in each group. For convenience we shall call them the Literary Group and the Scientific Group. The former was composed of persons from the various departments of the University which might be termed *literary* in contradistinction to those definitely *scientific*, and included one librarian, two professors of History, one professor and two associate professors of English, one professor and one assistant professor of Latin, two professors, two assistant professors, one instructor and one assistant in Romance languages, one assistant professor of Semitic languages, one professor and one assistant professor in journalism.

The I.Q.'s for this group ranged from 104.7 to 121.8, the average being 114.3, with 47 per cent. above and 53 per cent. below the average. Average time required for the test was 41 minutes, with 47 per cent. completing it in less than the average time for the group and 51 per cent. requiring more than the average time. Reference to Table III. will show the number of failures on each test. Of the five Verbal Tests only one was missed and this by 3 persons; with the exception of II—2 results of the Constructional Tests are fair and results on the Immediate Memory Tests make the poorest showing.

Consider next the Scientific Group which was made up of one assistant professor of mechanical practice, one professor of steam and gas engineering, two associate professors, two instructors, one student assistant and one research fellow

in physics, three assistant professors in chemistry, one assistant professor in agricultural chemistry, one assistant professor in agronomy, one instructor in manual arts, one associate professor of mathematics, one professor and one assistant professor of genetics. I.Q.'s here ranged from 107.3 to 121.8, the average was 113.7 and 53 per cent. were above while 47 per cent. were below this average. Time required varied from 27 to 60 minutes with 38½ minutes as the average and 59 per cent. requiring less than the average time for the group while 41 per cent. required more than the average time. Table III. shows the number of failures made in the Scientific Group on each test. Of the five Constructional Tests one is perfect and one was missed only once,

TABLE III

SHOWING THE NUMBER AND PERCENTAGE OF FAILURES ON EACH TEST BY EACH GROUP AND CLASSIFYING THE TESTS AS VERBAL, CONSTRUCTIONAL AND IMMEDIATE MEMORY TESTS. LIT. = LITERARY GROUP, SCI. = SCIENTIFIC GROUP, V = VERBAL, C = CONSTRUCTIONAL, I M = IMMEDIATE MEMORY

Test	No. Failures		Per Cent. Failures		Classification of Test
	Lit.	Sci.	Lit.	Sci.	
I—1.....	0	0	0	0	V
I—2.....	0	0	0	0	V
I—3.....	0	2	0	6	V
I—4.....	5	8	15	24	C
I—5.....	9	8	26	24	I M
I—6.....	3	5	9	15	C
Al—1.....	10	9	29	26	I M
Al—2.....	2	0	6	0	C
II—1.....	0	0	0	0	V
II—2.....	8	5	24	15	C
II—3.....	6	9	18	26	I M
II—4.....	3	2	9	6	V
II—5.....	9	8	26	24	I M
II—6.....	2	1	6	3	C

the showing on the other three being hardly so good as that on the Verbal Tests. As was the case with the Literary Group the Immediate Memory Tests are very poor.

Making now a somewhat more direct comparison of the results for the two groups, we find a number of facts of

interest. First, as regards the attitude shown by the two groups, on the whole the Scientific Group made better subjects, were less opinionated, adopted a more impersonal attitude toward the tests and seemed more interested in the purposes and results of the experiment. There were fewer objections to tests in general and fewer criticisms of individual tests as 'inadequate,' 'tiresome,' 'unfair,' 'no test at all,' than in the Literary Group. There was also less annoyance evidenced when a test was not passed. Of course the experimenter followed the technique recommended by Terman and the results were not made known to the subjects, but in such a test as the vocabulary test and in the logical memory test and the digit-span tests the subject was pretty well aware of the character of his performance without being told.

With regard to I.Q.'s there is a difference of only 0.6 between the average I.Q.'s for the two groups, not a sufficient difference to bear any significance. As to the time required to complete the tests the Scientific Group were apparently somewhat speedier, their average being 38½ minutes, while that for the Literary Group was 41 minutes. Only five, or 30 per cent. of the Literary Group, completed the tests within the average time of the Scientific Group. One reason for this time difference may lie in the fact that more definitions were given in the vocabulary test, and more exhaustive explanations of the difference between abstract terms, on the part of the Literary Group than on the part of the Scientific Group. Another reason may lie in the fact already mentioned, that there were fewer comments made in the nature of objections or criticisms during the progress of the test by the Scientific Group than by the Literary Group.

If, now, we classify the tests according to Miss Downey's method and indicate the number of failures on each test for each group we have as a result Table IV.

We find, taking all the Verbal Tests into consideration, that there are 3 per cent. fewer failures on the part of the Literary Group than on the part of the Scientific Group, while with the Constructional Tests exactly the opposite is true and the Scientific Group makes 3 per cent. fewer failures

TABLE IV

CLASSIFYING THE TESTS AS VERBAL, CONSTRUCTIONAL AND IMMEDIATE MEMORY TESTS AND SHOWING THE NUMBER OF FAILURES IN EACH TEST BY EACH GROUP

Verbal Tests

	I—1		I—2		I—3		II—1		II—4	
	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.
No.....	0	0	0	0	0	2	0	0	3	2
Per Cent..	0	0	0	0	0	6	0	0	9	6

Constructional Tests

	I—4		I—6		Al—2		II—2		II—6	
	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.
No.....	5	8	3	5	2	0	8	5	2	1
Per Cent..	15	24	9	15	6	0	24	15	6	3

Immediate Memory Tests

	I—5		Al—1		II—3		II—5	
	Lit.	Sci.	Lit.	Sci.	Lit.	Sci.	Sci.	Lit.
No.....	9	8	10	9	6	9	9	8
Per Cent.....	26	24	29	26	18	26	26	24

than the Literary Group. In the Immediate Memory Tests taken as a whole, there is a 1 per cent. difference in favor of the Scientific Group. So far, there seems very little evidence that the tests really enable differentiation between literary and scientific ability.

Let us, however, pursue the matter a little further and examine the results more closely by characterizing the performance of each subject on each test as *excellent*, *medium* or *poor*. Such a characterization of performance is necessarily based upon an arbitrary decision as to just what is meant by these standards, but since the same limits are set for both groups of subjects no unfairness can result from such a comparison. Arbitrarily, then, in the Verbal Tests

we set the following standards, P standing for *poor*, M for *medium* and E for *excellent:*

Test I—1	Score 69 or below	= P
	Score 70 to 84	= M
	Score 85 to 100	= E
Test I—2	Score below 8	= P
	Score of 8	= M
	Score of 10	= E
Test I—3	Less than 3 contrasts	= P
	Three contrasts	= M
	Four contrasts	= E
Test II—1	Score 79 or below	= P
	Score 80 to 89	= M
	Score 90 to 100	= E
Test II—4	Neither passage	= P
	One passage	= M
	Both passages	= E

In the Constructional Tests our standards are as follows:

Test I—4	Less than 3 correct	= P
	Three correct	= E
	Four correct	= E
Test I—6	Two errors—More than 4 minutes	= P
	One error—3′ 1″ to 4 minutes	= M
	One error—3 minutes or less or Two errors—2 minutes or less	= E
Test II—2	No division made. Either E or P, meaning in this case failure.	
Test II—6	One problem	= P
	Two problems	= M
	Three problems	= E

The Immediate Memory Tests we divide as follows:

Test I—5	One correct	= P
	Two correct	= M
	Three correct	= E

Test Al—1	None correct	= P
	One correct	= M
	Two correct	= E
Test II—3	One correct	= P
	Two correct	= M
	Three correct	= E
Test II—5	One correct	= P
	Two correct	= M
	Three correct	= E

Analysis of results according to these standards of performance gives the figures shown in Table V. In the Verbal Tests we have the following averages: 12.2 of the responses of the Literary Group were excellent as against 9 of the

TABLE V.

SHOWING THE NUMBER OF *Excellents*, *Mediums* AND *Poors* FOR EACH GROUP ON EACH TEST, TESTS BEING CLASSIFIED AS VERBAL, CONSTRUCTIONAL, ETC.
E = EXCELLENT, M = MEDIUM, P = POOR

Verbal Tests	Lit.			Sci.		
	E	M	P	E	M	P
I—1	17	0	0	13	4	0
I—2	10	7	0	12	5	0
I—3	14	3	0	10	6	1
II—1	14	3	0	6	9	2
II—4	6	8	3	4	11	2
Averages	12.2	4.2	0.6	9	7	1
Percentages	71.8	24.7	3.5	53	41.2	5.8
Constructional Tests						
I—4	11	1	5	7	2	8
I—6	8	4	5	7	1	9
Al—2	5	10	2	12	5	0
II—2	5	—	12	12	—	5
II—6	7	8	2	7	9	1
Averages	7.2	4.75	5.2	9	3.4	4.75
Percentages	42.3	27	30.7	53	20	27
Immediate Memory Tests						
I—5	2	3	12	1	3	13
Al—1	2	6	9	2	7	8
II—3	1	5	11	0	3	4
II—5	2	2	13	0	4	13
Averages	1.75	4	11.25	0.75	4.25	9.5
Percentages	10.3	23.5	66.2	4.4	25	70.6

Scientific Group; there were 4.2 responses classed as medium in the Literary Group and 7 in the Scientific Group; 0.6 of the Literary Group's responses were poor as against 1 of the Scientific Group. Or, putting it another way we may say that in the Literary Group we find 61 responses out of a possible 85 graded *excellent*, or 71.8 per cent., 24.7 per cent. *medium* and 3.5 per cent. *poor*, while in the Scientific Group we find 53 per cent. excellent, 41.2 per cent. medium and 5.8 per cent. poor. Judged in this manner the Literary Group does show itself superior to the Scientific Group on the Verbal Tests.

For those tests termed *Constructional* we find an average of 7.2 excellent responses for the Literary Group *versus* 9 for the Scientific Group; the Literary Group averages 4.75 medium responses, the Scientific Group 3.4; there are 5.2 poor responses in the Literary Group compared to 4.75 in the Scientific Group. Stated in terms of per cent., the Literary Group furnished 42.3 per cent. excellent, 27 medium and 30.7 poor responses as against the 53 per cent. excellent, 20 per cent. medium and 27 per cent. poor responses of the Scientific Group. Here the advantage is slightly in favor of the Scientific Group, though it is not nearly so marked as is the better showing of the Literary Group in the Verbal Tests.

The Immediate Memory Tests show so slight a difference between groups as to make the working out of percentages scarcely worth while. However, we find the Literary Group showing 10.3 per cent. excellent, 23.5 per cent. medium and 66.2 per cent. poor responses, while the Scientific Group shows 4.4 per cent. excellent, 25 per cent. medium and 70.6 per cent. poor.

There is still another way of comparing these two groups, that of further analyzing some of the test results. The tests which lend themselves most readily to a more complete analysis than we have given are the Vocabulary, Code and Ingenuity Tests. If we look more closely at the results of the Vocabulary Test (I—1 and II—1) we find in the Literary Group a score ranging from 86 to 100 and time required

from 2 to 8 minutes, averages being 95.3 for the score and 4′ 52″ for time required. It is in this test that we should expect to find the most marked superiority for the Literary Group and this seems to be the case. Scores for the Scientific Group range from 78 to 96, time required from 3 to 11½ minutes, with averages of 82 for the score and 6′ 12″ for the time. The Literary Group has an advantage, then, of 13.3 in score and 1′ 20″ in time in tests I—1 and II—1, both Verbal Tests.

Next, let us consider the results of the Code Test, I—6. Here we find errors for the Literary Group ranging from 0 to 8 with an average of 1.3. Time range is 1′ to 9′ with an average of 3′ 16″. The Scientific Group does not improve much upon these results, having an error range of 0 to 5, average 1.5, and a time range of 1½′ to 6′, average 3′ 30″. So we cannot grant any superiority in performance in this Constructive test to the Scientific Group.

Turning to another of the Constructional Group, the Ingenuity Test, II—6, and making a detailed analysis of its results we find '*a*' missed by one member of the Literary Group, but by none of the Scientific Group. Time range was from 20″ to 5′ for the Literary Group, average time 1′ 34″, while for the Scientific Group it was from 20″ to 4′ with an average time of 1′ 24″; '*b*' was missed by 6 of the Literary Group which required from 25″ to 5′ or an average of 2′ 9″ for this part of the test. Four of the Scientific Group failed on it, their time range being from 45″ to 5′, average time 2′. On '*c*' 7 of the Literary Group failed, time required was from 1 to 5 minutes, average time 2′ 42″, and 6 of the Scientific Group registered failure, time ranged from 45″ to 2½′ with an average of 1′ 26″. Combining the three parts of the test and taking the average for the whole we find the Literary Group requiring 6′ 25″ with 14 failures while the Scientific Group required 4′ 50″ and registered 10 failures, a difference in time of 1′ 35″ and in failures of 4. In this one of the Constructional Tests then, there does seem to be superiority on the part of the Scientific Group.

Conclusions.—Since none of the remaining tests are capable of so detailed an analysis, we shall not attempt further minute comparison, but will try to summarize briefly the results arrived at.

1. The first object of the research was "to determine the possible limits of performance for adults especially chosen because of their reputed brilliancy or genius." As was explained above, this purpose could not be completely carried out because of the difficulty of securing the desired subjects. Fifteen of those who took the tests may, however, be said to have the reputation in university circles of possessing exceptional ability. Tabulation of the I.Q.'s and the tests missed by these subjects appears below. (Table VI.) It

TABLE VI

SHOWING THE I.Q.'S AND TESTS MISSED FOR 15 SUBJECTS CLASSED AS EXCEPTIONALLY BRILLIANT

I.Q.	Tests Missed
121.8	
121.8	Al—1
119.2	I—5
118.7	II—2
118.7	II—2
118.7	II—4
118.7	Al—1, II—2
116.6	I—4, I—5
116.1	I—6, II—2
114.0	I—4, I—5, I—6
113.0	I—5, II—2, II—5
113.0	I—4, Al—1, II—3, II—5
113.0	I—5, II—3, II—5
110.4	I—4, I—5, Al—1, II—3, II—5
109.8	I—4, Al—1, II—2, II—3, II—5

will be seen that in one instance the I.Q. falls slightly below that of the Superior Adult level, that in two cases it reaches the highest possible score for an adult, I.Q. 121.8. It was a matter of surprise to the writer that only one subject made an absolutely perfect record. True, failure on either Al—1 or Al—2 had no effect on the test score or the I.Q., but it was her expectation that more subjects would have an I.Q. of 121.8 and that more would have perfect records. The fact that one subject not classed as 'brilliant,' whose record therefore is not included with those of the fifteen

mentioned, had an I.Q. of 121.8 and that several other reagents not included in the 'exceptional' list had higher I.Q.'s than some who do appear in it, leads again to a questioning of the real value of the I.Q. as now computed for adults. Certainly it is of no great help in close comparisons within a group such as we have been attempting here.

2. Recording the percentage of failures in each test as has been done in Tables I. and II. shows that, as was the case with college students, there are relatively too many failures on the Immediate Memory Tests to make them of real value for adults as they stand. It would be interesting to know whether visual presentation of these tests would result in fewer failures.

3. There is slight evidence that there is a difference in performance between the two groups, Literary and Scientific, on the two sorts of tests. This difference is more marked upon minute analysis of certain separate tests. It is not sufficiently great, however, to warrant the use of these tests as aids in assigning courses to college entrants, or in determining the 'bent' of any particular subject, unless supplemented by other tests.

4. One or two interesting problems growing out of this one might be mentioned as fields for future investigation. *a.*—Presentation of half of the Immediate Memory Tests visually and half auditorially, and comparison of the number of failures resulting under the two methods. *b.*—The large percentage of failures on I—4 and II—2, both of which depend so largely upon the imagination, suggest the giving of these tests to (1) a group of adults of college age who lack college training, (2) a group beyond college age who have no college training, (3) a college student group whose ages average about the same as those of group (1), (4) a faculty group corresponding in age and number to group (2). Comparison of results so obtained might show two things. First, the extent to which age affects the element of imagination; second, the extent to which training and education affect it. It might then be determined just how large a part tests of imagination *per se* should play in a test for General Intelligence.

CUMULATIVE CORRELATION

BY J. CROSBY CHAPMAN

Yale University

If we consider a variable x_1 to be dependent upon a number of other variables x_2, x_3, x_4, x_5, etc., we may weight, by the ordinary method of partial correlation, the values x_2, x_3, x_4, etc., so that the residual $x_{1.\,2.\,3\,\ldots\,n}$ will be a minimum. Under these conditions the partial regression coefficient of x_3, for example, is dependent not only on x_2 but also on x_4, x_5, x_6, etc. While in many cases this is what the conditions of the problem demand, there are certain practical situations in which the following question is pertinent. Having decided to use a variable number of tests in a serial order and weighting the first test unity, what weight must be attached to the succeeding test in order that the two tests combined will exhibit the maximum correlation? Furthermore, supposing the first and second tests are weighted according to the above procedure, and regarding this combination as a single variable, what must be the weight attached to another variable so that the new cumulative correlation is a maximum. At each step, in addition to the weighting, the resulting correlation is desired.

To translate into symbols, what is the value of m, n, p, etc., in the following correlations so that each correlation in succession may be a maximum: $r_{1.2+m3}$; $r_{1.2+m3+n4}$; $r_{1.2+m3+n4+p5}$; etc., where m is the weight attached to the third variable; n, that to the fourth; p, that to the fifth, etc. This procedure may be called cumulative correlation as opposed to straight partial correlation. It must not be confused with multiple correlation. In order to avoid misunderstanding, it is well before proceeding further to make clear what meaning the subscripts carry. Here $r_{1.2+m3+n4}$ is the correlation between x_1 and $x_2 + mx_3 + nx_4$ where the value of n is derived on the assumption that $x_2 + mx_3$ is a single

variable. It might, except for inconvenience, be written $r_{1.(2+m3)+n4}$. Similarly $\sigma_{2+m3+n4}$, which is used later is the equivalent of $\sigma_{(2+m3)+n4}$. The equation for five variables expressing the relation between x_1, x_2, x_3, x_4, x_5, is:

$$cx_1 = x_2 + mx_3 + nx_4 + px_5,$$

where c = constant defined later.

The formulæ for such operations are well known and comparatively simple. They are deduced from first principles below, though certain steps in algebra and calculus have been omitted. Let x_1, x_2, x_3, x_4, x_5, represent the variables, expressed as deviations from the means; the problem is what weight m must be attached to x_3 to make the correlation between x_1 and $x_{2+}mx_3$ a maximum. Writing this correlation as $r_{1.2+m3}$ we have

$$r_{1.2+m3} = \frac{\Sigma x_1(x_2 + mx_3)}{(\Sigma x_1^2)^{\frac{1}{2}}(\Sigma(x_2 + m_3)^2x)^{\frac{1}{2}}} \cdot$$

$$= \frac{r_{12}\sigma_2 + r_{13}\sigma_3 m}{(\sigma_2^2 + 2r_{23}\sigma_2\sigma_3 m + \sigma_3^2 m^2)^{\frac{1}{2}}} \cdot \qquad (x)$$

By differentiating with regard to m, and equating to zero, it can readily be shown that the value of m which makes $r_{1.2+m3}$ a maximum is given by the equation

$$m = \left(\frac{\sigma_2}{\sigma_3}\right)\frac{r_{13} - r_{12}r_{23}}{r_{12} - r_{13}r_{23}}. \qquad (a)$$

As might be expected, post facto, $m = b_{13.2}/b_{12.3}$ where, as in the usual Yule notation,

$$x_1 = b_{12.3}x_2 + b_{13.2}x_3.$$

Equation x may for convenience be written

$$r_{1.2+m3} = \frac{r_{12}\sigma_2 + r_{13}\sigma_3 m}{\sigma_{2+m3}}, \qquad (b)$$

where σ_{2+m3}, the denominator of equation x, is given by the formula

$$\sigma_{2+m3} = (\sigma_2^2 + 2r_{23}\sigma_2\sigma_3 m + \sigma_3^2 m^2)^{\frac{1}{2}}. \qquad (c)$$

We may now consider a and b as the basic formulæ; from

them the value of m, the most favorable weighting, and the value of the new cumulative correlation is obtained.

As the next step it will be necessary to make the following substitution

$$r_{4.2+m3} = \frac{r_{42}\sigma_2 + r_{43}\sigma_3 m}{\sigma_{2+m3}}. \qquad (d)$$

To add a further variable x_4, we consider the sum $x_2 + mx_3$ as a single variable, and by employing the basic formulæ a and b, proceed at once to write down the value of the weight n, to be attached to x_4 and also the value of the new cumulative correlation.

$$n = \left(\frac{\sigma_{2+m3}}{\sigma_4}\right)\frac{r_{14} - r_{1.2+m3}r_{4.2+m3}}{r_{1.2+m3} - r_{14}r_{4.2+m3}}. \qquad (e)$$

With this value of n determined

$$r_{1.2+m3+n4} = \frac{r_{1.2+m3}\sigma_{2+m3} + r_{14}\sigma_4 n}{\sigma_{2+m3+n4}}, \qquad (f)$$

where

$$\sigma_{2+m3+n4} = (\sigma_{2+m3}{}^2 + 2r_{4.2+m3}\sigma_{2+m3}\sigma_4 n + n^2\sigma_4{}^2)^{\frac{1}{2}}. \qquad (g)$$

For the next step the two following substitutions must be made:

$$r_{5.2+m3} = \frac{r_{52}\sigma_2 + r_{53}\sigma_3 m}{\sigma_{2+m3}}, \qquad (h)$$

$$r_{5.2+m3+n4} = \frac{r_{5.2+m3}\sigma_{2+m3} + r_{54}\sigma_4 n}{\sigma_{2+m3+n4}}. \qquad (i)$$

Proceeding to introduce x_5 with its most favorable weighting p, we have

$$p = \left(\frac{\sigma_{2+m3+n4}}{\sigma_5}\right)\frac{r_{15} - r_{1.2+m3+n4}r_{5.2+m3+n4}}{r_{1.2+m3+n4} - r_{15}r_{5.2+m3+n4}}. \qquad (j)$$

Whence, knowing p,

$$r_{1.2+m3+n4+p5} = \frac{r_{1.2+m3+n4}\sigma_{2+m3+n4} + r_{15}\sigma_5 p}{\sigma_{2+m3+n4+p5}}, \qquad (k)$$

where

$$\sigma_{2+m3+n4+p5} = (\sigma_{2+m3+n4}{}^2 + 2r_{5.2+m3+n4}\sigma_{2+m3+n4}\sigma_5 p + p^2\sigma_5{}^2)^{\frac{1}{2}}. \qquad (l)$$

This process of combining each added variable with all

the previous variables weighted as shown, and substituting in equations *a* and *b* may be carried on indefinitely. At each stage, a new weighting is derived, and the increase in correlation shows the value derived from the addition.

It is perhaps advisable to point out specifically what are the differences between the weightings as obtained by this method [1] and the ordinary regression coefficients. It has already been mentioned that $m = b_{13.2}/b_{12.3}$, so that m is so chosen that x_1 and $x_2 + mx_3$ yield a maximum correlation. No other factor entered into the determination of m; its value is absolutely *independent* of the variables x_4, x_5, x_6, etc. The value of the corresponding partial regression coefficient is *dependent* on all the other variables x_4, x_5, x_6, etc. For the accurate investigation of independent causation the partial regression coefficients are the correct instruments; but for the investigation of the effect of adding a single test to a battery already weighted as though complete in itself, the cumulative correlation process has a distinct field of usefulness because it provides a weighting for any variable introduced in a definite series which is independent of the inclusion or exclusion of any of the later variables in the series. More important is the fact that it may also be employed, with a small amount of error, for the practical purpose of determining the weights of a given series of tests. The values of the weightings, so obtained, will not differ very considerably from those found by the ordinary partial procedure.

For purposes of illustration, four problems[2] have been selected for which the usual partial regression coefficients and $R_{1.(234\text{—}n)}$ were available. For each of these problems the partly analogous values found by the cumulative method have been determined by continued substitution in equations

[1] I understand from correspondence with Dr. H. A. Toops that he is using a somewhat similar procedure. Not having had an opportunity to examine the equations he employs, I cannot tell the extent to which his assumptions are identical with those made in this article.

[2] Problem 1 is taken from Yule's 'Theory of Statistics,' p. 242. Fourth Edition. Problems 2 and 3 were kindly supplied me by Dr. J. E. Anderson who had already determined the partial regression coefficients. Problem 4 is taken from Kelley's 'Educational Guidance,' p. 94.

a to *g* for the four-variable problem, and by using equations *a* to *l* for the five-variable problem. The variables are included in the order shown by the subscript; as a general rule, this follows the magnitudes of the zero order correlation coefficients with the variable x_1. As will be readily seen, for each order of the variables chosen, there will be corresponding alterations in the values of the resulting weightings *m*, *n*, *p*, etc. The best plan in ordinary work is to introduce the independent variables in the order of the magnitudes of their zero order correlations with the dependent variable.

Problem I. $r_{12} = .52$, $r_{23} = .49$, $\sigma_2 = 41.7$,
$r_{13} = .41$, $r_{24} = .23$, $\sigma_3 = 5.5$,
$r_{14} = -.14$, $r_{34} = .25$, $\sigma_4 = 23.8$.

By ordinary partial correlation procedure

$$c'x_1 = 1.0x_2 + 4.3x_3 - 1.2x_4,$$

where $c' = (b_{12.34})^{-1}$ (Yule notation).

$$R_{1.234} = .63.$$

By cumulative correlation procedure

$$cx_1 = 1.0x_2 + 3.7x_3 - 1.1x_4,$$
$$r_{1.2+m3+n4} = .62.$$

Problem II. $r_{12} = .33$, $r_{23} = .46$, $r_{35} = .46$, $\sigma_2 = 6.6$,
$r_{13} = .29$, $r_{24} = .39$, $r_{45} = .47$, $\sigma_3 = 5.7$,
$r_{14} = .25$, $r_{25} = .49$, $\sigma_4 = 5.8$,
$r_{15} = .23$, $r_{34} = .29$, $\sigma_5 = 7.3$.

By ordinary partial correlation procedure

$$c'x_1 = 1.00x_2 + .86x_3 + .72x_4 + .00x_5,$$
$$R_{1.2345} = .39.$$

By cumulative correlation procedure

$$cx_1 = 1.00x_2 + .81x_3 + .64x_4 + .00x_5,$$
$$r_{1.2+m3+n4+p5} = .38.$$

Problem III. $r_{12} = .33$, $r_{23} = .46$, $r_{35} = .29$, $\sigma_2 = 6.6$,
$r_{13} = .29$, $r_{24} = .25$, $r_{45} = .20$, $\sigma_3 = 5.7$,
$r_{14} = .27$, $r_{25} = .39$, $\sigma_4 = 5.4$,
$r_{15} = .25$, $r_{34} = .31$, $\sigma_5 = 5.8$.

By ordinary partial correlation procedure

$$c'x_1 = 1.00x_2 + .70x_3 + 1.05x_4 + .63x_5,$$
$$R_{1.2345} = .41.$$

By cumulative correlation procedure

$$cx_1 = 1.00x_2 + .81x_3 + .98x_4 + .65x_5,$$
$$r_{1.2+m3+n4+p5} = .41.$$

Problem IV. $r_{12} = .72$, $r_{23}, = .61$, $r_{35} = .55$, $\sigma_2 = 5.19$
$r_{13} = .62$, $r_{24}, = .61$, $r_{45} = .59$, $\sigma_3 = 5.17$ } assumed
$r_{14} = .58$, $r_{25}, = .82$, $\sigma_4 = 5.14$ } equal.
$r_{15} = .63$, $r_{34}, = .66$, $\sigma_5 = 5.19$

By ordinary partial correlation procedure

$$c'x_1 = 1.00x_2 + .50x_3 + .24x_4 + .12x_5,$$
$$R_{1.2345} = .76.$$

By cumulative correlation procedure

$$cx_1 = 1.00x_2 + .53x_3 + .24x_4 + .13x_5,$$
$$r_{1.2+m3+n4+p5} = .76.$$

If it is desired to pass from the relative weightings of the variables to prediction values and the absolute weightings in this procedure, it is necessary to determine the value of the constant c, for each problem. From the fact that $x_2 + mx_3 + nx_4 + px_5$ is considered as a single variable, it follows that in case of five variables

$$\frac{\mathrm{I}}{c} = \frac{r_{1.2+m3+n4+p5}\sigma_1}{\sigma_{2+m3+n4+p5}}.$$

Finding the value of c for each problem and substituting we get for the cumulative equations

Problem I. $x_1 = .33x_2 + 1.23x_3 - .38x_4.$
II. $x_1 = .33x_2 + .26x_3 + .21x_4 + .00x_5.$
III. $x_1 = .29x_2 + .23x_3 + .28x_4 + .19x_5.$
IV. $x_1 = .37x_2 + .20x_3 + .088x_4 + .047x_5.$

These may be compared with the corresponding partial equations

Problem I. $x_1 = .33x_2 + 1.38x_3 - .38x_4.$

II. $x_1 = .32x_2 + .27x_3 + .23x_4 + .00x_5.$

III. $x_1 = .30x_2 + .21x_3 + .31x_4 + .19x_5.$

IV. $x_1 = .36x_2 + .18x_3 + .086x_4 + .047x_4.$

It is also interesting to trace the rise in value of the cumulative correlation coefficient with the addition of each successive variable. This is shown below:

Cumulative Correlation Coefficient	Problem I	II	III	IV
r_{12}	.52	.33	.33	.72
$r_{1.2+m3}$	.55	.37	.37	.75
$r_{1.2+m3+n4}$	.62	.38	.40	.76
$r_{1.2+m3+n4+p5}$	—	.38	.41	.76

From the statements that have been made and the results that have been shown, it can be seen that for the practical purpose of weighting the variables in accordance with their contribution to the total effect, there is no great discrepancy, up to five variables, between the ordinary partial correlation procedure and this much simpler cumulative method employing equations, the logic of which is fairly apparent, and which are relatively few in number compared with those necessary in the usual partial correlation work.

LEARNING WHEN FREQUENCY AND RECENCY FACTORS ARE NEGATIVE

BY JOSEPH PETERSON

George Peabody College for Teachers

Introduction

Of the more specific factors operating within the larger situation that compels learning, frequency effects seem to be principally, if not entirely, responsible for the backward elimination of errors in the maze. Such error elimination, and such operation of frequency factors, have been most clearly brought out in mental maze experiments, which eliminate the influence of various disturbing and irregular spatial factors. It has been shown that in mental-maze learning the probability of one's getting successfully past a blind alley without subsequent returns is greater near the goal-end than at the entrance-end, and that the learning coefficient constantly increases from the entrance to the goal.[1] In a ten-blind maze the learning coefficient at the first blind was found to be 1.031, whereas it was 1.333 at the tenth, for the first trial. This advantage at the latter part of the maze for the elimination of entrance to blinds seems to be based on frequency factors, and is traceable to the fact that in each trial the forward runs will necessarily exceed the backward runs by unity. This may be indicated by the equation

$$f - 1 = b,$$

in which unity is the only constant term. The absolute values of f and b constantly decrease toward the distal part of the maze, and therefore the ratio of f to b increases, being greatest at the last blind. This is what gives the frequency

[1] Peterson, Joseph, 'The Backward Elimination of Errors in Mental Maze Learning,' Jour. Exper. Psychol., 1920, 3, 257–280. The 'learning coefficient' of a blind was defined in this report as 'the ratio of the expected number of runs past that blind in the forward direction *to* the expected number of runs into it (from either direction).' These expectations are based on the laws of probability.

advantage to the forward runs past blinds, over the backward runs and runs into blinds, an advantage that increases regularly toward the distal end of the maze.

If frequency factors could be exactly balanced against themselves so far as positive and negative effects are concerned, and if the same thing could be done with recency factors, we should have a rather unique situation for the study of learning, and could try out experimentally some of the current views regarding the nature of the learning process, at any rate in a particular if not in a crucial case. To attempt to do this was the object of the present investigation.

It will be necessary on the start to distinguish between frequency of stimulus and frequency of response. The one sort of frequency effects may be balanced without a balance of the others being effected. In the present study we have attempted to balance exactly the frequency-of-stimulus factors making, respectively, for and against learning at any part of the maze, and we have also attempted to do this for the recency factors. At the same time, when the effects of frequency and of recency of *response* are considered, we find that both in the case of frequency and in that of recency the odds are greatly against learning; that is to say, these factors are negative in their effects, tending to fix the erroneous responses rather than the correct ones. In a later study we hope exactly to equate both the stimulus and the response frequency and recency factors.

It may seem confusing to distinguish between the operation of frequency and of recency factors in the stimulus, on the one hand, and in the response on the other, but the distinction seems necessary if we clear our study of certain confusions which may pass unnoticed in mere theoretical discussions but which are met when exact experimental control is attempted. In the former case, frequency of stimulus, the stimuli making for changes in behavior, are probably entirely extero-ceptive, while they are both proprio-ceptive and extero-ceptive in the latter. However, an analysis on this basis is not yet practicable. When we speak of response in the present report, we must consider a choice as a unitary re-

sponse, and must not for the present attempt to analyze the responses further. We are attempting to get a quantitative hold on our problem, and to do this it is necessary that we thus define our terms. This meaning of response is, however, not far different from that currently used in discussions of learning factors, except that our units are more nearly equal, being always simply a choice of one of a presented pair of letters. We do not deny that behind different choices lie different backgrounds, but these differences are nevertheless produced by the choices themselves which we have attempted to make equal.

Method of Procedure

The method of the present investigation is identical, so far as the experimental part is concerned, with that used in our earlier study cited above, with the important exception that on each wrong choice the subject is immediately brought back to the entrance and again given the alternative choices at the first blind, without any indication being made to him that anything irregular has happened, and without anything occurring to call his attention to the fact that a return has been made. Of the various maze situations that have been presented either to animals or to man, such a procedure is of course possible only with the mental maze. The instructions to the subject are these:

A maze, you know, is a winding way to some goal, but it has many blind alleys which will lead you to error if you enter them. I have drawn a maze which I will show you when you have completed this experiment. Wherever there is a choice of two alleys, each pathway is designated by a letter. No two parts are lettered the same. I call out two letters at a time, and you are to choose one of them; then I call out two more; and so on, till you get to the goal. Whether the right letter comes first or second in the pair called out, is determined wholly by a chance-order schedule, arranged by flipping a coin; so you need not attend to which letter comes first. In fact, if you try to choose according to some pre-determined plan or order, or try to make the letters you choose spell words, you will never learn the way to the goal. The problem is to see with how few errors you can get to the goal, and how soon you can learn to make no errors at all. You are through the experiment when you get to the goal the third successive time without any error. You will be told each time when the goal is reached, and also the number of errors made in reaching it. Remember, accent and order of calling out the letters have no significance. Don't attend to them. Keep in mind also that where we go in the maze depends on your own choices.

Note.—Please do not inform any one else about the nature of this experiment. All subjects must begin ignorant of it.

Let us illustrate the process of carrying out the experiment by following an actual record. The experimenter having Figure 1 before him and sitting behind a screen so that unconscious facial expressions and movements will not influence the subject, hands the instruction sheet to the latter, asking him to read it carefully. While the subject is reading the instructions the experimenter records his name, the date, and any other conditions of the experiment to be noted; and just when the subject returns the sheet the experimenter also records the exact time to the nearest minute, and immediately calls out *N-V*. The subject chooses *N*, which the experimenter records and underlines, and he also puts a cross over it because the choice is an error. The subject, of course, does not see the record. *V-N* are then called out in the same tone of voice and accent as before, the right alternative now being called first as indicated in the chance-order schedule. This time *V* is chosen. It is quickly recorded and underlined, and *I-L* are presented. *I* is chosen, and is re-

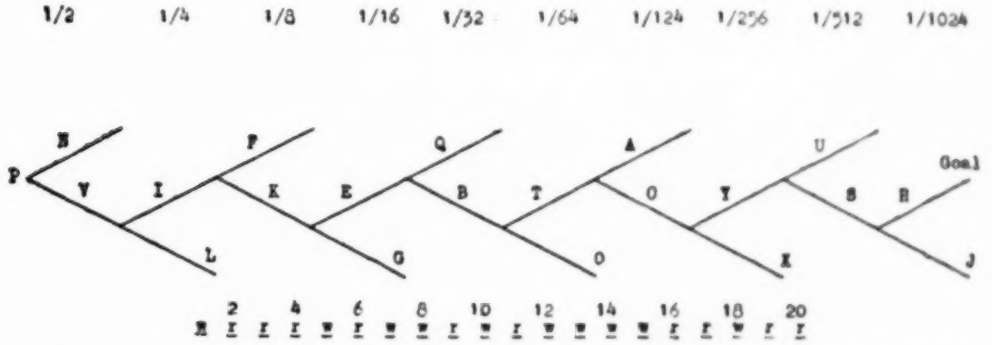

FIG. 1. The mental maze and the chance-order schedule. The fractions above the maze indicate the chance of passing in a single forward run the corresponding right alternatives, or of entering the wrong alternatives.

corded and underlined, after which *K-F* are called out. *F* is now chosen and is recorded by the experimenter with a cross over it. *N-V* are called, *N* coming first because *w* is the fifth letter in the schedule, which the experimenter has been following with his left index finger, and this is the fifth choice. Each *w* in the schedule means to call out the wrong letter

first, while each *r* indicates that for the stimulus it represents, the right alternative must be presented first in the pair. *V* is now chosen by the subject and is recorded without any qualifying mark whatever, because it is neither the first of the alternatives called nor a wrong choice. Thus the process continues, probably for nearly an hour, when the subject finally reaches the goal. The experimenter now says "Goal!" and after recording the time, quickly counts and announces the number of errors. He then tells the subject that the second trial begins, and noting that the second letter in the schedule is an *r* (for right first), he calls out *V-N*. Each trial thus begins as indicated by the letter of its number in the chance-order schedule so that the order of presentation in successive trials is never the same, even if the subject makes no errors, or if he makes the same errors in two different trials. It will be noted that in each case in which the subject chooses the first of the two alternatives presented, the choice is recorded and underlined, and, in addition, erroneous choices are marked with a cross so that they can readily be counted up at the close of the trial. The subject when he reaches the goal is told the number of errors made but he must himself find out which of his many choices are wrong.

When the experiment is completed the experimenter has a full record of all the stimuli presented to and the responses made by the subject, even to showing at each alternative whether the first or the second of the letters presented at any time was selected. He can therefore study in detail any subjective tendency noted in the choices, whether it be to choose the first or the second of the alternatives in each pair, or to prefer now the one and now the other in any larger regular or irregular cycle. A thorough study of the subject's choices may in any case be helpful in throwing light on the subject's learning and on his peculiar difficulties. On the other hand, the subject himself has had a real problem to solve, one on which he has a good deal of room to exercise his own constructive ability and to try out various hypotheses, reactions which are often obvious enough to the experimenter.

This learning process is therefore not a matter of merely

remembering a series of syllables, words, or objects in a fixed order, determined in advance by the conditions of the experiment, as has usually been the case in association experiments. And it is important to note that in the present form of learning, the subject, in the solution of his problem, by whatever methods he may adopt, leaves a complete, permanent record that can be analyzed in detail, even to determining the laws of association influential in each choice. From such a record, then, we are able to check up on some of the theories which flourish so well in the absence of objective facts.

A glance at the maze, Fig. 1, will show that the chances are even, that is, 1 to 1, that the subject will make the right choice at the first blind; that they are 1 to 4 that he will also successfully pass the second blind; 1 to 8 that he will get past the first three blinds; and so on. The chances of getting past all the blinds without being thrown back is, of course, the product of all the chances of passing the several blinds by themselves, or $\frac{1}{2}$ raised to the power of the total number of blinds. In the present maze the chances of reaching the goal without any error whatever is $(\frac{1}{2})^{10}$, or 1 to 1024. None of my subjects has yet succeeded in doing this, although some of them have had very good luck. Now if any one of the blinds is chosen and an immediate return to the beginning is made by the experimenter's calling out the *N-V* alternatives again, it is obvious that the mere frequency and recency of the wrong choice will operate toward the continuation of the error at this juncture, that is, *against* the learning of the maze. Indeed, the chances are that frequency and recency response effects will thus in 1023 times out of 1024 count negatively for the learning, while their stimulus effects are equally balanced for and against learning. That the effects of frequency and recency factors in response are very real is obvious to anyone who examines in detail the individual records as tabulated for intensive study. We have here then a case of constructive learning in which frequency and recency influences are decidedly against the formation of the proper associations for the solution of the problem.

Subjects

Twenty college students and graduate students were used as subjects in the present experiment with the modified mental maze, five men and fifteen women. None of them had had previous experience with either form of the mental maze. No introspections were called for during the experiment, as that would introduce variable factors, but some of the subjects made significant remarks, either spontaneously during the experiment or in reply to questions after it was completed. These remarks became part of the record, as did also various notes on the behavior of subjects during the experiment. Four of the subjects, all women, were unable to learn the maze in the time available, so their records are not included in this report. Three of these subjects worked 72, 98, and 110 minutes respectively, without any promise of success, while the fourth one, a woman who was having difficulty in carrying some of her courses, persisted in the invariable repetition of her erroneous choice of *N*, declaring that she was 'all in the dark anyway, and might as well go one way as another.' Our results are therefore based on the records of sixteen subjects, all of whom completed the problem. In every case the learning was completed at a single sitting.

Results

Table I. shows the general distribution of errors made by sixteen subjects to the point of complete learning of the maze. All but seven of the subjects made no errors after the first trial, but this trial was sometimes very long. The errors made in all trials after the first, by seven subjects, are shown separately in the table; and the total number of entrances to the goal, or goal choices, is also indicated, because these choices must be taken into consideration in the determination of the expected choices in the several parts of the maze. The table shows the distribution of errors of each subject separately as well as the total number of errors at each bifurcation by all the subjects, first for the first trial only, and then for all subsequent trials by seven subjects. It also gives the

number of errors (based on the total number of errors made *plus* the goal choices) that would be expected at each part of the maze on the law of probability. On the basis of probability alone we should expect as many errors in the first blind

TABLE I

DISTRIBUTION OF ERRORS IN THE SEVERAL PARTS OF THE MAZE BY SIXTEEN SUBJECTS

Subject	Parts of Maze										Total Errors	Goal Choices	Total
	H	*L*	*F*	*G*	*Q*	*O*	*A*	*X*	*U*	*J*			
First trial													
A.....	10	4	3	2	1		1	1	2	1	25	1	26
B.....	85	49	11	7	3	3	1				159	1	160
C.....	31	6	2	4							43	1	44
D.....	41	28	15	1	36	1	1	1		1	125	1	126
E.....	61	19	15	3	6		4	11		1	120	1	121
F.....	112	21	17	8	4	3	2	1			168	1	169
G.....	88	55	9	7	5	1		1			166	1	167
H.....	130	41	36	21	29	9	2	1	1	1	271	1	272
I.....	2	1	2	4	2	2	1	1	1		16	1	17
J.....	24	52	16	5	23	3	2		1		126	1	127
K.....	142	42	22	6	7	11	1	1	1		233	1	234
L.....	12	10	5	9	3		1	2	2		44	1	45
M.....	6	1	2	1	1	1		1	1	1	15	1	16
N.....	22	16	5	6	3	5	4	1	1	1	64	1	65
O.....	13	9	5	4	2	1	1	1	1		37	1	38
P.....	141	69	17	11	4	3	4	1	1	2	253	1	254
Total....	920	423	182	99	129	43	25	24	12	8	1,865	16	1,881
Expected.	940.5	470.3	235.1	117.6	58.8	29.4	14.7	7.3	3.7	1.8	1,879.2	1.8	1,881
All subsequent trials													
A.....									1	1	2	1	3
B.....			1	1	3		3	3			11	1	12
C.....	3	32	30	6	1	4	2	2			80	5 (2–6th)	85
E.....									1		1	1	2
G.....										1	1	1	2
I.....				1							1	1	2
L.....	1										1	1	2
Total, all.	924	455	213	107	133	47	30	29	14	10	1,962	27	1,989
Expected.	994.5	497.3	248.6	124.3	62.2	31.1	15.5	7.8	3.9	1.9	1,987.1	1.9	1,989

as in all the following blinds *plus* the number of goal choices. This is because the chances are just even for choosing the first blind or going past it, and all the passes must end in some later blind or in the goal. The same thing holds for each subsequent blind; the choices made at the juncture where

it appears should be divided evenly between it and all blinds beyond it, with the goal choices included in the latter. It will be seen that of the first-trial choices there were 920 errors made at the first blind, while 940.5 were to be expected; at the second blind there were 423 errors and 470.3 expected; and so on, a constantly decreasing series both in actual and in expected errors toward the distal end of the maze. Table II. shows these same totals in per cent., the expected errors, given in bold type, being the same in the first as in the subsequent trials or in all trials combined.

TABLE II

PERCENTAGE DISTRIBUTION OF ERRORS IN THE MENTAL MAZE (SIXTEEN SUBJECTS)

	Parts of Maze										Total Errors	Goal Choices	Total
	N	*L*	*F*	*G*	*Q*	*O*	*A*	*X*	*U*	*J*			
First trial, Obt'd..	48.9	22.5	9.7	5.2	6.8	2.3	1.3	1.3	.6	.4	99.0	.9	99.9
All trials, Obt'd ..	46.5	22.9	10.7	5.4	6.7	2.4	1.5	1.5	.7	.6	98.9	1.4	100.3
Expected.....	**50.0**	**25.0**	**12.5**	**6.3**	**3.1**	**1.6**	**.8**	**.4**	**.2**	**.1**	**100.0**	**.1**	**100.1**

In general the correspondence of actual to expected errors in all parts of the maze is rather close, both in the first trials only and in all trials together of all the sixteen subjects who completed the maze learning. This shows that probability laws are indeed operative to a large extent even though the choices are voluntarily made by the subjects. But in both the first trials and in all trials combined we find a larger percentage of errors toward the distal end of the maze than expected on the laws of chance. This tendency of increased errors beyond expectations toward the goal is, of course, the effect of learning, whatever has brought about the learning, an effect that is usually very obvious to the experimenter during the first trial. Many subjects, in fact, make no errors after once reaching the goal, having learned to avoid every blind before this place is reached.

A very striking effect of the modification in this experiment effecting complete returns after erroneous choices is that which shows itself in the order of elimination of errors. *In*

the present form of the mental maze errors are eliminated in the forward direction, whereas in the previous form not involving a return with each error, as well as in the mazes used for animals, the opposite tendency appears, as we have already seen. It was shown that backward elimination was mathematically attributable to the operation of frequency factors, which, as we mentioned above, are in the present case actually negative. In both cases, however, frequency and recency factors are not the real determining conditions of the learning process. This is the important result of our studies that will receive more explicit attention in a later part of this report.

The forward elimination of errors in the present form of the mental maze is revealed in a study of Tables III. and IV.

TABLE III

DISTRIBUTION OF THE FINAL ERRORS OF THE SEVERAL SIXTEEN SUBJECTS IN THE MODIFIED MENTAL MAZE

	Parts of the Maze										Total
	N	*L*	*F*	*G*	*Q*	*O*	*A*	*X*	*U*	*J*	
Number of final errors.....	2	0	1	1	0	1	1	2	4	4	16

Table III. shows the number of *last errors* by our sixteen subjects in each blind of the maze. All but five of these subjects made their last errors at or beyond the seventh blind, and only two made their final errors in the first blind, one of these stating after the experiment that his error at this place was due to a choice purposely made to verify an idea that he had respecting it. Half of the sixteen subjects made their final errors in the last two blinds.

But the tendency to eliminate errors in the forward direction in the modified mental maze is even more obvious when we follow an individual in detail in the process of the learning. In Table IV. the progressive elimination forwards is shown by merely numbering all the errors of any one subject in the order of their occurrence, and by putting these several

numbers in the columns corresponding to the blinds in which they occurred. The table shows separately the distribution and the order of occurrence of the errors made by three subjects, *E*, *G*, and *M*. Space does not permit the giving of the responses of some of the longest records, but those given are

TABLE IV

THE DISTRIBUTION AND ORDER OF ALL THE ERRORS MADE BY CERTAIN SUBJECTS IN THE MODIFIED MENTAL MAZE

	Parts of Maze									
N	*L*	*F*	*G*	*Q*	*O*	*A*	*X*	*U*	*J*	
Subject *E*										
1–9	10–13									
14–21	22									
23–26	27									
28–33	34									
35–40	44, 46	41, 42		43, 45						
48–50		51		47						
52–54	55									
56–59				60						
64	61–63									
68, 69	65, 67		66							
72, 73	70, 76	74, 75		71						
77–79	80, 81									
82, 84, 85				83		86				
87, 88		90, 91	89							
92		94	93			95–97	99		98	
107, 108		100, 101					102–106			
		109, 110					111			
		112					113			
114, 115		116, 117					118, 120			Goal
Second trial								121		Goal
Subject *G*										
2–58, 64	1, 59, 60	62	61	63						
66, 73–74	65, 70–72	67–69								
76, 78, 82	75, 77, 80						79			
84–86, 92	81, 83, 89	87, 88	90							
93, 99, 101	91, 95, 96	94, 97		98, 100						
104, 106	102, 103									
110, 111	107–109									
113, 115	112, 114									
124	116–123									
143, 146	125–142	144								
148, 150	145, 147									
152, 153	149, 151		154, 156							
155, 158	157		159							
160, 161			163, 165	162, 164	166					Goal
Subject *M*										
2, 3, 5, 6	1	4	7							
8, 9		12		10	11		13	14	15	Goal

fairly representative, the first two of the medium or poor students, and the last, the responses of subject *M*, of the best subjects. The table is to be read as follows: Subject *E* chose *N* of the first alternative nine times in succession; the next four times he passed the first blind successfully, that is, chose *V*, and in every case entered the second blind. Then again he entered the first blind successively eight times. The twenty-second error was made at the second blind after having successfully passed the first. Errors 23 to 26, inclusive, are again in the first blind (at the first alternative), number 34 is in the second, 35 to 40 in the first, *N*, 41 and 42 are in the third, or the *F* blind, the first two having been correctly passed; and so on. Note that for each error in any blind beyond the first one, the subject must successfully choose the right alternative at all previous junctures or critical points; for example, error number 85 was made in the first blind, then the next choices at the first six blinds were all correct, and the 86th error was made by entering the seventh blind, *A*. Both subject *E* and subject *G* learned very slowly to avoid the early blinds, and they were never quite sure of passing them successfully till near the end of the experiment. On the contrary, subject *M* readily learned to avoid choices that were unfruitful, the result being a more regular forward progression of errors, with the last error occurring in the final blind, *J*.

In general, the method of learning may be characterized as follows: One usually begins by making a great number of errors in two or three of the first blinds, according to what would be expected on the laws of probability. These errors are repeated, sometimes a surprising number of times, especially by the passive, rather conventional type of person. Variations in response then occur, often without any purpose or forethought, and the way becomes cleared to the blinds beyond, which in turn give trouble just as the first ones did. In correcting the errors at these alternatives the subject often loses the effect of his earlier responses and again falls into error at the blinds first encountered. These are more readily eliminated now, however, and the subject gradually pushes his conquest over errors further and further back toward the distal end of the maze. In proportion as the subject keeps in

mind the situation, or retains the effects of his earlier experiences, the series of repetitive responses each terminating in an error become more and more far-reaching until finally the goal is reached. In later trials few errors, if any, are made. That is, the subject as a rule clears his way as he goes, gradually enlarging his range of orientation, until the whole situation is either consciously reacted to as a unit, or at any rate is made effective somehow in the determination of his reactions. Some subjects seem rather suddenly to come to a recognition of the means of locating their errors and hence eliminate them, and it is usually obvious to the experimenter just when and at what point this recognition comes. It is really surprising, though, how slowly this idea, that erroneous choices throw one back to the *N-V* alternative, dawns on some of the subjects. In fact, the poorer subjects have constantly to recur to the first errors—the choice of *N*, *L*, or *F*—and they rather gradually and unconsciously wear themselves out, as it were, of these erroneous responses. Their responses differ from those of the better subjects in the fact that in the enlargement of the range of reaction already described, they frequently lose what they have gained and lapse back into earlier erroneous responses terminating the shorter ranged reactions that are characteristic of the earlier stage of the learning. But when once the principle, that errors throw one back, is clearly perceived, these relapses do not occur. The difference between the good and the poorer subjects is principally one of the explicitness with which the situation and the principle of error elimination are grasped. But from our experiments with the mental maze, as well as from those described as rational learning,[1] it seems that one may come to follow a useful principle in learning even though one is not aware of the fact or of the principle. Presumably many persons in practical life are unconsciously forced by the general consistency of circumstances thus to follow principles, and it would be a great error to ascribe the uniformity of such behavior to imitation and suggestion as is commonly done by certain writers on social psychology.

[1] Peterson, Joseph, 'Experiments in Rational Learning,' *Psychol. Rev.*, *25*, 1918.

There seems, then, in the case of all subjects, good or poor, to be some gradually accumulative effect of the successive repetition of errors in a monotonous situation, an overlapping of stimulus effects of some sort that amounts to an increase in the range of receptivity to the stimuli of a complex situation, some effect the neural and general physiological basis of which we are yet very ignorant, but which reveals itself very well in the reactions and the verbal statements of the subjects of the present experiment. For instance, one bright, mentally alert boy in Peabody Demonstration School, on whom the experiment was tried, soon noted the return to the *N-V* situation in connection with certain of his choices, and he laughed heartily each time when he was thrown back. Using that return as a criterion of the degree of success of his choices and explorations in the maze, he soon solved the problem. Several of our subjects noted unpleasant affective states associated with the recurrence of the first often repeated letters, *N-V*, *I-L*, etc. One adult woman, who had been thrown back a great number of times, finally showed signs of attention to the recurrence of the *N-V* situation, and, after the completion of the problem, said that the letters at these early blinds got to be very repugnant to her. "I just got to hate these letters," was her way of expressing it. Some subjects after long repetitions without any sign of recognition of the fact that errors threw them back, would in time perceptibly sigh when the *N-V* situation was presented, following an error; and yet after such evidence of discomfort they frequently continued to make the error, ignorant of the principle which might help them reach the goal. It is thus clear that something more than cerebral activities became invloved in connection with these repeated failures, physiological processes which are undoubtedly in large part visceral. The means by which these accumulative organic effects arise and inhibit the erroneous responses in learning is unknown.

We have already indicated that a number of our subjects perceived clearly, after a few experiences of being thrown back, the significance of these set-backs; and that they checked their errors explicitly and voluntarily by them. Such

subjects made much more rapid progress in the elimination of errors at successively encountered blinds, and they seldom made any lapses into errors once overcome, as did subjects who could not at the end of the experiment explain how to overcome errors. But we must not overlook the fact here that this more explicit perception is itself only the conscious aspect of an intra-organic process now possibly more completely cerebral in its location than is true when strong affective disturbances, not accompanied by clear perception of the meaning of the situation, arise. To deny 'stamping in' effects of right responses or 'stamping out' effects of wrong responses to feeling states *per se* and to ascribe such effects to perception or to the recognition of meanings, is of course inconsistent. In both cases there are physiological bases of which we are yet ignorant, the ignorance being in part due to the fact that we have been content to stop with the mental states themselves in our tracing of causal relations. This error is not unlike that of attributing certain acts to will, to attention, or to any other of the spontaneous faculties still in good repute in much of our current psychological literature. The perception of the meaning of the set-backs is itself but a response, a kind of behavior revealed in the reactions of the subject, even though expressed verbally by the subject.

An Analysis of the Influence of Frequency and Recency Factors

Having considered the method of learning in the modified mental maze and the order of elimination of errors, we may now turn our attention to a consideration of the rôle of frequency and recency factors in the process. It will be recalled that during the learning by each subject a complete record was made of all the successive alternatives presented to him as well as one of his individual responses in the exact order of their occurrence. From this record we are, fortunately, able to tabulate each subject's responses in detail, both as to order and location in the maze (or in the series of stimulus situations) and as to their agreement or disagreement with

expectations based on frequency and recency factors as determined by all past responses up to the point in question. The first step in this analysis is to tabulate in a form like that shown in Table V., all the responses of each subject. A separate line is allowed for each forward movement from the initial alternatives at *N-V*, and on each of these lines are placed pencil checks in the appropriate columns for the several choices, each check indicating a choice of the letter (or part of the maze) at the head of the column. Each forward movement continues, of course, until some error is made, or until the goal is reached. When all of the subject's responses to the point of no further errors are tabulated in this manner, the record is ready for an analysis of the frequency and recency laws of response. For this purpose we use the following abbreviations: *B* indicates a response that is in accordance with the expectations based on both recency and frequency, while *b* is for a response of the opposite nature, contrary to frequency and recency expectations; *Rf* represents a response in agreement with recency expectations and contrary to frequency expectations, and *Fr* is just the reverse of *Rf; R* is for a response in agreement with expectations on recency when frequency factors just balance, while *r* is just the reverse, contrary to recency; finally 1 designates the first response to any one of the alternatives at a blind, a response which cannot therefore be analyzed according to recency and frequency factors. With these symbols in mind we go over the individual record and write in ink over the erasable checks the appropriate symbol. When the ink dries the checks are erased, leaving only the symbols. We get then a record such as that shown in Table V. of subject *L*. A study of this record will make the process clear. *L* first made three erroneous choices of *N* at the first alternative. The first one is marked 1, and the other two, being in agreement with both recency and frequency expectations, are marked *B*. It is obvious that now if these laws determine the learning all subsequent responses should be choices of *N;* but the fourth response is a choice of *V*, contrary to both these laws, hence this response is marked *b*. The subject now having passed

Table V

An Analysis of All the Responses of Subject *L* to the Point of Complete Learning of the Mental Maze. *B* means that Response is in Agreement with Expectations based on Frequency and Recency; *b*, the Reverse, Contrary to these Laws; *Fr*, in Agreement with Expectations on Frequency, Contrary to Recency; *Rf*, the Reverse of *Fr*; *R*, in Accord with Recency when Frequency Factors Balance; *r*, Contrary to Recency; 1, first Response to a Pair of New Alternatives

Trial	Parts of the Maze																			
	N	*V*	*L*	*I*	*F*	*K*	*G*	*E*	*Q*	*B*	*O*	*T*	*A*	*C*	*X*	*Y*	*U*	*S*	*J*	*H*
(1) ...	1																			
	B																			
	B																			
		b	1																	
		Rf	*B*																	
		Rf	*B*																	
		R	*B*																	
		B		*b*	1															
	b																			
		Fr		*Rf*	*B*															
	b																			
		Fr		*Rf*	*B*															
		B		*Rf*	*B*															
		B	*r*																	
		B	*B*																	
		B	*B*																	
		B	*B*																	
		B	*B*																	
	b																			
	Rf																			
	Rf																			
	Rf																			
		Fr		*b*		*b*		1	1											
						—														
	b																			
		Fr		*Rf*		*Rf*	*b*													
		B	*Fr*																	
	b																			
		Fr		*b*		*Rf*		*r*	*B*											
				—																
		B		*Rf*		*Rf*	*b*													
		B		*Rf*		*R*	*R*													
	b																			
		Fr		*Rf*		*B*	*B*													
		B		*R*		*B*	*B*													
		B		*B*		*B*	*B*													
		B		*B*		*B*	*B*													
		B		*B*		*B*	*B*													
		B		*B*		*B*		*b*	*B*											
								—												
		B		*B*		*B*		*Rf*	*B*											
		B		*B*		*B*		*Rf*		*b*	1		1							
										—	—									
		B		*B*		*B*		*Rf*		*Rf*	*B*		*b*			1		1		
													—							
		B		*B*		*B*		*Rf*		*Rf*	*B*		*R*		*B*		*B*			
		B		*B*		*B*		*Rf*		*Rf*	*B*		*B*	*b*						
		B		*B*		*B*		*R*		*R*	*B*		*B*		*Rf*					
		B		*B*		*B*		*B*		*B*	*B*		*B*		*r*		*b*		1	
															—		—		—	
(2) ...	*b*																			
		Fr		*B*		*B*		*B*		*B*	*B*		*B*		*B*		*Rf*		*B*	

Summary of L's Reactions

Trial	Law	*N*	*V*	*L*	*I*	*F*	*K*	*G*	*E*	*Q*	*B*	*O*	*T*	*A*	*C*	*X*	*Y*	*U*	*S*	*J*	*H*	Total	Wrong	Right
1st.	*B*	2	22	7	11	3	13	5	1	3	1		5		3		1	1				78	21	57
	R		1		1		1	1	1		1				1							7	1	6
	Rf	3	2		7		3		5		3					1						24	4	20
	Fr		6	1																		7	1	6
	b	6	1		3		1	2	1		1				1	1			1			18	9	9
	r			1					1								1					3	1	2
2d..	*B*				1		1		1		1		1		1		1				1	8	0	8
	R																					0	0	0
	Rf																		1			1	0	1
	Fr		1																			1	0	1
	b	1																				1	1	0
	r																					0	0	0

the first alternative (the first blind) chooses *L* of the *L-I* alternative. This being the first choice here, is marked 1. Again *L* chooses *V*, which is in agreement with recency of response but contrary to frequency, and is so marked. So we go on through the entire record till the goal is reached. *L* made one error in the second trial, the choice of *N*, so all the second-trial reactions are also recorded. A summary of *L*'s reactions as here classified is given at the base of the table, but in this summary the uncritical reactions marked 1 are omitted. In this summary of the first trial we find, for example, a total of 78 *B*-responses, 21 of which were wrong and 57 of which were right choices; a total of 7 *R*-responses, with 1 wrong and 6 right. There are 24 *Rf*-responses with 4 wrong and 20 right; 7 *Fr*-responses, with 1 wrong and 6 right; 18 *b*-responses of which 9 are right and 9 wrong; and 3 *r*-responses, 1 wrong and 2 right.

In this way each of the sixteen records has been treated, and the several summaries have been condensed into Tables VI., VII., and VIII. The method is seen to be similar to that which we applied to the learning of white rats in the animal maze.[1] In Table VI. are summarized the various critical or significant responses of the sixteen subjects, sepa-

[1] Peterson, Joseph, 'Frequency and Recency Factors in Maze Learning by White Rats,' *Jour. Animal Behav.*, 1917, 7, 338–364.

TABLE VI

CLASSIFICATION OF ALL THE SIGNIFICANT RESPONSES MADE BY THE SIXTEEN SUBJECTS IN THE MODIFIED MENTAL MAZE

	Kind of Response						
	B	*R*	*Rf*	*Fr*	*b*	*r*	Total
First Trial							
Number right	1,104	78	508	203	445	29	2,367
Number wrong	839	43	232	279	288	27	1,708
Total	1,943	121	740	482	733	56	4,075
2d to 6th trials							
Number right	213	10	67	22	27	7	346
Number wrong	20	6	14	17	36	4	97
Total	233	16	81	39	63	11	443
All trials (summary)							
Number right	1,317	88	575	225	472	36	2,713
Number wrong	859	49	246	296	324	31	1,805
Total	2,176	137	821	521	796	67	4,518
All trials in per cent.							
Per cent. right	**29.1**	**1.9**	**12.7**	**5.0**	**10.4**	**.8**	**60**
Per cent. wrong	**19.0**	**1.1**	**5.4**	**6.6**	**7.1**	**.7**	**40**
Total	**48.1**	**3.0**	**18.1**	**11.6**	**17.5**	**1.5**	**100**

TABLE VII

CLASSIFICATION OF ALL THE SIGNIFICANT RESPONSES OF THE SIXTEEN SUBJECTS IN THE MODIFIED MENTAL MAZE TO THE POINT IN EACH PART OF THE MAZE WHERE THE LEARNING WAS COMPLETE

	Kind of Response						
	B	*R*	*Rf*	*Fr*	*b*	*r*	Total
Number right	702	38	437	225	472	36	1,910
Number wrong	859	49	246	296	324	31	1,805
Total number	1,561	87	683	521	796	67	3,715
Per cent. right	**18.9**	**1.0**	**11.8**	**6.1**	**12.4**	**1.0**	**51.4**
Per cent. wrong	**23.1**	**1.3**	**6.6**	**8.0**	**8.7**	**.8**	**48.6**
Total per cent.	**42.0**	**2.3**	**18.4**	**14.1**	**21.1**	**1.8**	**100.0**

TABLE VIII

PERCENTAGE OF RESPONSES OF SIXTEEN SUBJECTS IN AGREEMENT WITH EXPECTATIONS BASED ON FREQUENCY AND ON RECENCY FACTORS, AND ALSO CONTRARY TO THESE EXPECTATIONS

Responses are classified according to whether they are right or wrong.

	Kind of Response						Total		Difference
	B	*R*	*Rf*	*Fr*	*b*	*r*			
Right responses									
With frequency	18.9			6.1			25.0		
Against frequency			11.8		12.4			24.2	
Wrong responses									
With frequency	23.1			8.0			31.1		
Against frequency			6.6		8.7			15.3	
Totals	42.0		18.4	14.1	21.1		56.1	39.5	16.6
Right responses									
With recency	18.9	1.0	11.8				31.7		
Against recency				6.1	12.4	1.0		19.5	
Wrong responses									
With recency	23.1	1.3	6.6				31.0		
Against recency				8.0	8.7	.8		17.5	
Totals	42.0	2.3	18.4	14.1	21.1	1.8	62.7	37.0	25.7

rated first on the basis of first and later trials, and then of all trials together up to the point of complete learning. The summary of all trials together is also given in percents, this part of the table being in bold face type. In this table, are summarized the classifications of 4518 responses, exclusive of insignificant responses. It is obvious that the analysis of the first trial responses is more important for the understanding of learning than is that of later trials, after the goal has been reached, because in later trials the effect of the learning has already become a considerable amount at different parts of the maze. In the first trial there were 1104 *B*-responses (in harmony with recency and frequency) in the right choices and 839 in the wrong ones; but there were also 955 correct choices *contrary* to frequency expectations (*i.e.*, 508 *Rf* + 445 *b*) and only 520 (or 232 *Rf* + 288 *b*) wrong responses contrary to frequency. Moreover, in the right responses there are

78 + 508, or 586, in accordance with recency, and 445 + 29, or 474, against recency expectations. Of the wrong responses there are 275 (or 43 *R* and 232 *Rf*) in accordance with recency expectations and 315 (or 288 *b* and 27 *r*) against such expectations. Surely nothing very decided for frequency and recency influences can be made out from these data. A detailed study of individual records in the form of Table V. shows, indeed, the influences of frequency and recency, but these influences are of such a nature as to fix the erroneous responses as much as they fix the correct ones. In fact, as we have already said, learning in the present form of the maze problem must go on *in spite of* the effects of both frequency and recency. But we have previously shown, in the last reference cited, that even in the usual type of maze learning by rats, learning effects must in many cases go directly against frequency and recency influences.

But the data in Table VI. are somewhat deceiving in as much as certain parts of the maze are learned earlier than other parts, and hence in these places of earlier learning the responses will continue to be one hundred per cent. in agreement with both frequency and recency expectations until the other parts are also learned. Thus in Table V. we note that column I. has 12 *B*, 1 *R* and 3 *Rf* correct responses, and column *K* has 14 *B*, 1 *R*, and 3 *Rf* correct responses beyond the last correct choice. These are cases of over-learning in which the responses must of necessity be one hundred per cent. in accordance with either frequency or recency, or both. It is possible to eliminate all such over-learning responses from our tables and find the influence of frequency and recency factors only up to the point of learning, say one response correct after which no errors occur in each part of the maze. This has, therefore, been done in the present study by omitting from our summaries all the responses occurring in each individual's record at places where learning is complete, counting only the first of the final series of correct responses at each place. For example, in case of subject *L* whose record is shown in Table V. all the responses below the horizontal lines in the several columns under right choices

have been omitted. The right choices are *V*, *I*, *K*, *E*, *B*, *T*, *C*, *Y*, *S*, and *H*. The results of this method are summarized in Table VII., in which we find, therefore, the total number of each kind of response up to the point of complete learning in the several parts of the maze brought together. These results are shown both in absolute numbers and in per cents. in the same table. Here we do not need to take note of the several trials, but can consider all trials together. The summary of per cents. affords easy comparison with the results given in Table VI. obtained by the more unsatisfactory method. In Table VII. it is seen that expectations based on frequency and recency of responses are satisfied oftener in the wrong responses than in the right ones. This is probably due to the persistence of wrong responses when once started in any part of the maze, a persistence that had to be overcome by factors other than those here studied.

Table VIII. is an analysis of the percentage data given in Table VII. While this analytical table shows a preponderance of responses in accordance with frequency and recency on the whole, it also shows that in the case of frequency there are more wrong responses in accordance with the law than there are right responses, and more right responses against it than wrong responses against it. In the case of recency there is almost an exact equality of *R*- and *r*-responses in the right choices on the one hand, and in the wrong ones on the other; that is to say, *agreement with recency has no relation at all to the rightness of the response, and in the case of frequency, agreement with the law is somewhat more likely to occur in the wrong than in the right responses while disagreement is more likely to be found in the right than in the wrong responses.* This result agrees with our theoretical determinations in the early part of this paper, that learning in the present form of the maze is not effected by recency or frequency factors, or by both combined, and it gives additional support to our conclusion of the year 1917, based on the learning of white rats in the ordinary maze problem, that such learning must go on in spite of—rather than because of—frequency and recency influences. This statement ap-

plies only to the *modification of the order* of successive responses and the inhibition of irrelevant acts, not, of course, to the mere fixing of any act, whether right or wrong, in a determined series. So far as the writer is aware, his is the first case of experimental work that has put to actual quantitative test the current and orthodox doctrine, coming down to us from the extreme association psychology, that learning is brought about wholly or in large part by the effects of recency and frequency factors. We have now on hand extensive results of experiments on about six hundred white and negro children in multiple choice learning, which agree substantially with our present results in affording evidence against this doctrine derived from the study of 'the association of ideas.'

Comparison with the Results of Other Experiments and Theoretical Considerations

The data of the earlier experiment on maze learning were obtained from an analysis of the responses of seventeen white rats in their first three trials in mazes. Two mazes of different form were used, but there is close correspondence between the results in the two cases. We shall therefore give only the summary table here for easy comparison with our present results.

Table IX

A Summary of the Distribution of 1,232 Responses in the First Three Trails of Seventeen White Rats in Maze

Figures give the per cent. of total responses and symbols have same meaning as in Table V.

Trial Number	*B*	*R*	*Rf*	*Fr*	*b*	*r*	Totals
1	33.7	8.6	2.9	3.3	34.9	16.5	100
2	41.4	6.9	5.4	9.1	27.2	9.9	100
3	48.3	6.1	6.3	9.2	23.4	6.6	100
1–3 together	41.7	7.1	5.1	7.7	27.9	10.5	100

In this table the symbols have the same significance as in the present study. Unfortunately, in the case of animal learning, no account was taken of whether the reaction in each case was right or wrong, but the complete study, to

which the reader is referred, shows no evidence that frequency and recency factors bear any real causal relation to the learning. Indeed, though it was not possible in the animal maze to effect complete returns with each error, as in the present experiment, evidence showed that frequency and recency tended more strongly toward the fixing of wrong acts than toward the correction of errors, and that responses in accordance with the expectations based on these laws increased in successive trials (as a *result* of the learning), reaching one hundred per cent. with perfect learning.

While no other quantitative study of frequency and recency effects in constructive learning seems to have been made, there has recently appeared from the Chicago laboratory an interesting and ingenious study of a complex form of learning based on ideation and controlled by quantitative methods.[1] Unfortunately the author, although paying some degree of attention to frequency factors, or repetition,—without always being clear as to what it is that is repeated—accepts the traditional view of the effects of frequency, and makes no quantitative analysis with a view to evaluating such effects. It is, of course, possible that his method and data do not permit of such an analysis. Numerous statements as to the effects of repetition are made, however, statements that are not supported by adequate evidence, if by evidence at all; and they seem to be contradicted by other statements. The methods and results are too complex to explain here; we can take note only of some of the statements regarding the method of learning. We are told that "Perhaps the most important factor in the abstraction and selection of the essential elements of our problems is frequency of repetition of responses to those elements . . . those elements which served most frequently as stimuli were the first to acquire definite meanings and special value for purposes of control" (p. 104). Again: "The most obvious factors in the selection and accentuation of essential elements were frequency of repetition of elements and their relative

[1] Peterson, John C., 'The Higher Mental Processes in Learning,' *Psychol. Monog.*, 920, *18*, No. 7 (Whole No. 129).

nearness to a goal, or end of action" (p. 120). Yet explanations in subjective terms are used to supplement these effects of frequency and nearness to the goal, and certain mental functions seem to be elevated almost to the position of faculties with spontaneous powers, a practice that is still current in much of our psychology and that is, indeed, hard to avoid wholly, in appearance at least, because of the complexity of mental processes and of their stimulation by various indirect means. Thus we find in the present case that the all-important frequency factors are supplemented by 'attention,' by 'foresight of consequences,' by 'perception of relations,' by 'the setting up of hypotheses,' etc. In the early stages of learning, "owing to a lack of acquaintance with the elements of the problem, the span of attention to the elements is very narrowly limited. Subjects who have passed beyond the early stages of learning also revert to this form under the stress of emotion, self-consciousness, or fatigue, which tend seriously to narrow the span of attention to the elements of the problem. There are some individuals who, through native limitations of the span of attention seem doomed to a marked dependence upon this primitive form of arousal of hypotheses. They drew rapidly and, disregarding recently observed and often well-known facts, frequently jumped to unwarranted conclusions from which they showed little ability to extricate themselves by crucial tests" (p. 112). "Sometimes hypotheses originated in explicit analysis and comparison of the elements of one or more series. . . . At its best this process involves a clear insight into the causal relations of the phenomena in question and carries its own confirmation" (p. 113). The nearness of any response to a goal is also emphasized, as in such cases, as well as in less complicated circumstances, the consequences of tentative or hypothetical draws can be foreseen (47 ff.).

The *facts* of behavior here described agree well with those that have been observed and mentioned in our own experiments on the higher forms of learning; but the explanations not only imply spontaneous mental forces, but also controls within which frequency must operate. This implication con-

tradicts the author's expressed view as to the importance of this factor. Occasionally these mental powers seem to be built up as a direct result of repetition, so that we seem to move in a circle. Indeed the author unwittingly reveals the fact that frequency factors operate only within limits imposed by certain other factors which really give direction to the learning that is either strengthened or counteracted by frequency and by recency. These quotations remind one of Thorndike's well-known explanation of selection in learning in terms of frequency, based on the very curious assumption that the various elements in the situation are 'responded to each once,' [1] an assumption of some controlling factor that prevents the mere repetition of the first erroneous act, expected on the basis of his all-sufficient principle of frequency as well as on that of recency. The sailing is, of course, clear when this assumption is granted.

Attention, foresight of ends and of consequences, hypotheses, etc., are but the subjective aspects, as it seems to me, of a selective process that is itself to be explained, and the explanation of it—though it certainly lies within the relationships of the life processes and the total stimulating conditions—is undoubtedly one of the most important problems before psychology today. The study in question, though it overlooks this problem, concealing it inadvertently behind subjective terms or mental forces, nevertheless gives many observations which seem to me to point out the direction of a solution. The author not only notes, as our own researches have shown, that subjects frequently get to respond selectively on the basis of some principle before they become explicitly aware of the principle,[2] but he also recognizes 'the combined associative pull of a number of disconnected elements observed in rapid succession' (pp. 111, 112). The explicit perception of relations and of principles seems to be but the mature manifestations, in subjective terms, of selective responses based upon the cumulative results of successive and simultaneous stimuli, not the causes of them. The

[1] Thorndike, E. L., 'The Psychology of Learning,' 1913, p. 41.

[2] *Op. cit.*, pp. 53, 109.

several quotations given above, as well as the objective results of the experiments, are entirely compatible with such a view. It is therefore probable that a detailed study of the responses in this type of complex analytical learning would also have revealed, as our present study has done, the inadequacy of frequency and recency factors as directors of the learning process.

Our notes as well as our quantitative data show that in the modified mental maze learning, the effects of any stimulus and of its resulting response do not fade away immediately. The range of the series of responses, as we have seen, gradually enlarges itself by breaking away from such repetition of errors as the effects of frequency and recency would bring about. On the subjective side of the responses we noted, in correspondence with these changes, that certain choices come to be avoided either with a dislike for the stimuli immediately following them—the *N-V* situation—or with explicit recognition of the fact that they are errors and throw one back. A cumulative effect of the successive stimulus-reaction processes seemed to build itself up and to inhibit errors. The neural changes corresponding to this modification in behavior are not yet understood, but the general fact itself that some change goes on bringing about an overlapping of stimulus effects is indisputable.

When an individual is stimulated in any manner the nerve impulses flash through the organism in but a small fraction of a second. But there is considerable evidence to show that the effects of this stimulus do not so immediately fade away. Probably the responses of muscles and glands set up other afferent nerve impulses, which in turn, being reflected up through cortical and subcortical coördinating centers, or synapses, out through efferent nerves, bring about further responses; and so on. Thus streams of impulses must pour in and out in various directions for some time, gradually decreasing in intensity—or in extent according to the 'all-or-none' law—until the disturbance is entirely abated except for changes of permeability of synapses that it has effected. These streams of impulses therefore will exist con-

temporaneously with subsequent stimuli and exert important directive influences on the nerve impulses these stimuli set up. Thus certain cumulative dispositions to responses in a measure consistent with all the stimulating circumstances will gradually take shape in the organism. All stimuli, moreover, have an effect on the organism in proportion to the general organic needs at the time of their occurrence. Thus when one is in a state of hunger the food stimulus is heightened. Every response tends to be complete, or consummatory, to the extent that it brings about a satisfaction—in the biological and not necessarily the psychic sense of the term—of these organic needs. It may bring recognition of self by others, or food, or it may promote self assertion, or arouse sex activity, or stimulate play activity when there is readiness for it in the way of 'surplus energy' or what not, etc. The overlapping of stimulus effects may thus be conceived as due, in part at least, to the rebounding of nerve impulses which involve the action of visceral muscles and glands as well as of skeletal muscles, and to the interactions of various organs upon one another.

It is this conception which the writer outlined in 1916, doubtless very imperfectly, as an hypothesis susceptible of experimental test.[1] A theory that does not lend itself to some sort of verification by experiment is of little value. It has been found, in conformity with the expectations based on the theory in question, that rats learn more readily to avoid errors in a maze with short blind alleys than in one with long blind alleys but in every other respect exactly identical with it.[2] It is not, of course, contended that no other explanation of these facts is possible; none has yet been given, however, and it seems clear that frequency and

[1] Peterson, Joseph, 'The Completeness of Response as an Explanation Principle in Learning,' *Psychol. Rev.*, 1916, *23*, 153–162. This first statement of the theory seems to have been badly misunderstood by one reviewer, who, by an unfortunate omission, misquotes an important passage. See Wiltbank, R. T., 'The Principles of Serial and Complete Response as Applied to Learning,' *ibid.*, 1919, *26*, 277–286. Mr. Wiltbank, by some means that I cannot understand, concludes that my theory neglects present external stimuli, and he consequently sets up to explain certain data, a theory that seems to be identical with my own theory except in name.

[2] Peterson, Joseph, 'The Effect of Length of Blind Alleys on Maze Learning: An Experiment on Twenty-four White Rats,' *Behav. Mon.*, Ser. No. 15, 1917.

recency factors fall short here. The present study shows in another form of learning, with human subjects, the inadequacy of these factors to account for the modification in behavior effected.

Indeed, many writers have shown some degree of recognition of this inadequacy of frequency and recency factors as complete explanation principles in learning. Thorndike has explicitly argued against the assumption that learning can be based on frequency of exercise,[1] and has added his 'law of effect,'[2] which is, of course, not an explanation at all but only a very rough statement of fact. Watson, though generally taking a rather extreme view with reference to the importance of frequency factors in learning, has recognized the probable operation of such other 'physiological principles' as 'reënforcement, inhibition, and summation of stimuli.'[3]

In the general psychology texts we still find attention, will, the perception of meaning, etc., playing the rôle of real selective agents in higher forms of learning; hence the inadequacy of the traditional associationism as applied to learning has not been so evident to the uncritical reader. It is in animal psychology, where these 'faculties' have not so obviously functioned, or where the assumption of such *forces* has been more obviously questionable, that the real problem of learning has been laid bare. It is here that our ignorance of the process has become evident. In this field the first tendency—as is seen in the writings of Lloyd Morgan and Thorndike, among the experimentalists—was to invoke pleasure and pain as the selective agencies. On the basis of frequency and recency alone it has thus far been impossible to explain how an animal reacts to situations, as we know that we ourselves do, rather than to a mere succession of stimuli.

But even if we grant the simultaneous operation of many stimuli on an individual by the means suggested above—the overlapping of stimulus effects—we are still not free from the obligation to explain how the more consistent acts survive

[1] Thorndike, E. L., 'Animal Intelligence,' 269 ff.

[2] Thorndike, E. L., 'Psychology of Learning,' p. 4.

[3] Watson, J. B., 'Behavior,' 275 f.

over the less consistent ones, or otherwise stated, how erroneous acts are eliminated. While in its technical aspects this is fundamentally a problem for neuro-physiology, the suggestion here of an hypothesis subject to experimental test may be *a propos*. If we grant that the basic life processes, controlled by stimuli (mainly internal) through the autonomic nervous system, are not only continuous but in a very real sense consummatory, it may be agreed that other acts are subservient and contributory to them. The synapses involved in the neural impulses directly affecting these autonomic processes may be thought of as permeable to a high degree. When these processes become impeded in any manner nerve impulses flow into various channels bringing about random, uncoördinated acts involving the skeletal muscles. How are these impulses again drained off as it were into the channels of the consummatory acts? Let us assume that each of these uncoördinated acts, causing conflict in the animal's behavior, brings about a shower of afferent impulses from the skeletal muscles, and that the consummatory acts—including not only the autonomic acts but also the acquired acts or habits making for their realization—bring fewer of these rebounding, afferent impulses because of less conflict. This assumption is certainly supported by introspective as well as by objective evidence; and the consummatory acts seem to spread their effects less extensively throughout the organism than do the random, uncoördinated acts. Now, it can be shown mathematically, on exactly the same principles that we have applied to the backward elimination of errors in the maze,[1] that these random impulses will be eliminated in the order of their proximity, in neural relationship, to the consummatory acts. The learning coëfficient, based on frequency, will be greatest at the synapses most closely associated with the autonomic processes basic to the present consummatory acts, because of the greater scarcity of rebounding impulses from these activities. Or if there is no greater scarcity of afferent impulses from these activities, the impulses may not be reflected back so directly to the specific synapses involved in the random conflicting acts, but may tend only to heighten the consum-

[1] See the first reference, above.

matory processes themselves and thus add to the general motivation. It is well known that certain processes, such as play and sex activities, tend toward their own continuation. In either event it can be shown mathematically that the spreading impulses will gradually drain off into the channels of those promoting the consummatory acts, first at the synapses next to those of these end-acts, and then, when these synapses have been sufficiently changed in resistance, at those next to them.

We shall have to develop this hypothesis more fully in another paper. The important point here is that such an application of our theory of maze learning is also applicable to nerve impulses in learning; and that it can be tested out experimentally by a study of the order of elimination of random acts in relation to their proximity to the consummatory acts.

Inhibition and facilitation must also be reckoned with in any complete theory of learning, but we are yet a long way from understanding their respective rôles in the modification of behavior.

Our experimental results, by a quantitative method applied to the laws of frequency and recency in learning, seem conclusive against the adequacy of these factors as explanation principles in learning; and these results must stand or fall independently of any relation to the theoretical considerations in this paper. It is plain that our knowledge of the real neural basis of learning is yet very limited. Whatever the neural processes in detail prove to be, it would seem that intelligence in general, and as manifest in learning in particular, consists in an arrangement of neuro-muscular mechanisms of such a nature that the organism is not subject to the mere contingency of the order of the occurrence of stimuli, but reacts to situations on the basis of general consistency. The degree of complexity of any situation to which one can learn to react efficiently is dependent largely on one's innate ability. On the basis of such a view it is obvious that one may profit by errors as well as by successes, a fact that is not easy to explain on the frequency-recency view.